Impact of Mobile Speed Enforcement (MPE): An Analysis of the Duration between Collisions at Enforced Sites

Samaa Agina[1] and Karim El-Basyouny[2]*

[1]MSc. Student, Department of Civil and Environmental Engineering, University of Alberta

[2]Associate Professor, Department of Civil and Environmental Engineering, University of Alberta

*****_Corresponding author_

Address: 6-362 Donadeo Innovation Centre for Engineering, 9211 – 116th Street NW, Edmonton, Alberta, Canada T6G 1H9

Final Report

May 2022

• EXECUTIVE SUMMARY

This report is submitted to the City of Edmonton in partial fulfillment of deliverables for the research project titled "Impact of Mobile Speed Enforcement (*MPE*): An Analysis of the Duration between Collisions at Enforced Sites." The primary goal of this research is to study the potential impact of the *MPE* deployment efforts on the duration between two consequent collisions. The project involved studying 250, 175, 212, and 219 sites in 2019, 2018, 2017, and 2016, respectively. These sites had varying traffic volume levels, roadway categories, and conditions. The project was performed in two phases: preparing the data for testing and applying a rigorous statistical analysis.

When deployed systemically, *MPE* forces drivers to comply with the posted speed limit and adopt a safer driving behaviour. The main objective is to explore the influence of deployed *MPE* on the duration between collisions, particularly to enhance efficiency when deploying limited enforcement resources and improving traffic safety. The report will also cover survival analysis and hazard models in evaluating the City's *MPE* program.

The data was obtained from the City of Edmonton's Safe Mobility Section, containing collisions data, *MPE* data, and site information. This data was compiled into a single spreadsheet using Site ID information. Subsequently, we scripted a MATLAB code to tabulate the data to facilitate a survival analysis. Finally, we applied the survival analysis to the data and investigated the relationship between the *MPE* variables and the duration between collisions.

Our analysis supports the positive effect of the deployed *MPE* hours and visits on increasing the duration between consequent collisions, which correspondingly reduced the risk of collision occurrence. For instance, the ratio between deployed *MPE* hours to visits had the highest impact on reducing the risk of collision. Moreover, we demonstrate that the grouped sites with above-average MPE variables had a higher survival probability than those with below-average *MPE* variables.

Table of Contents

List of Tables

List of Figures

1 INTRODUCTION

Traffic collisions claim over 1.2 million lives and cause about twenty million injuries annually (*1*). In Canada, for example, at least five people die daily due to traffic collisions on roads. The road safety problem is even more profound in the US, with reports indicating that about 90 deaths occur daily (*1*). However, those figures are gradually reducing with governments adopting systemic, multi-approach intervention strategies (*2, 3*).

One approach has been to use Automated Speed Enforcement (*ASE*) programs, especially where speeding has been identified as a major cause of collisions. *ASE* has proven to reduce both the severity and frequency of collisions. Previous findings have shown *ASE* programs to reduce crashes by 8.9-to-51% while simultaneously reducing collision fatalities and injuries by 12-to-50% (*4*). Such programs have been effective in many Canadian cities like Winnipeg, Calgary, and Edmonton.

In Edmonton, data indicate that *ASE* programs have reduced collisions by up to 20% (*5*). The City of Edmonton's primary goal when employing these programs is to minimize traffic crashes and enhance traffic safety. Achieving that implies that the authorities must prioritize and map collision risk sites. Accordingly, they must give precedence to *ASE* deployment at high collision sites. Findings from Alberta suggest that Mobile Photo Enforcement (MPE) deployment effectively improves safety at these sites (*6*).

This study emphasizes the effectiveness of *MPE* use in reducing the duration between two consequent collisions, demonstrated through hazard models, achieved by analyzing collision severity, collision frequency, exposure of collision risks, traffic counts, and MPE deployment data.

Collision sites can be identified by establishing the Equivalent Property Damage Only (*EPDO*) for every kilometre over three years. This method transforms all collisions into an equivalent number of damage-only collisions by assigning heavier weights to the various collision severity (i.e., injuries, fatalities, and property damage) levels. This measure combines both collision severity and frequency to formulate a single performance criterion that incorporates crash counts with higher severity (*7*). Severe cases can be further normalized based on the road lengths by dividing the EPDO collision frequency by the total road kilometres. This measurement helps to establish and compare risks to exposed collision locations. Arguably, it is easier to collect road length data because several cities in Canada have authenticated databases that have such information.

Several national and international reports have shown that the introduction of *ASE* programs has had positive outcomes by improving traffic safety (*8*). For instance, reviews of findings worldwide between the late 1990s and 2000s show that the introduction of *MPE* effectively reduces speeding vehicles and collisions by 82 and 51 %, respectively (*9*). According to Li (*7*), the *MPE* program is a reliable tool to reduce crashes since the perceived risk of detection is heightened with each subsequent deployment, which results in an induced general deterrence leading to lower speed violations.

1.1 Background

Red-light running and speeding are among the leading causes of road collisions in the US and Canada (*4*). Research suggests that both red-light running and speeding increase one's risks of being involved in a crash and the risk of death and injuries. Studies have shown how these two infractions have varied consequences based on magnitude, which increases collision severity during such instances. Indeed, research has shown that the chances of being involved in a collision and speeding cases are directly proportional (*4*). According to Evans, when one increases the average speed by roughly 1%, the fatality risk rises between 4 and 12% (*10*). Simultaneously, doubling the speed doubles the risks of collisions, injuries, and death, while exceeding the average speed by 20% escalates the risks six times the norm (*10*). Previous meta-analyses engaging more than ninety-eight studies concluded that the interrelations between road safety and speed are substantial and meet all the categories of causality. Elvik (*11*) argues that speeding is the single highest determinant in traffic fatality cases. Moreover, there is a significant difference in fatality risk when a moving vehicle (speed dispersion) moves faster than the surrounding traffic (*11*).

Studies suggest that red-light running often results in right-angle collisions, believed to be more severe than other collision types (*12*). In the US, for example, red-light running crashes are estimated to cost more than US$14 billion (*13*). On the other hand, excessive speeding contributes to more than 18% of crashes (*13*). That translates to over 2,000 injuries and deaths every year. Meanwhile, red-light running accounts for more than one-quarter of the traffic injuries (*14*). Studies conducted by the Ministry of Transportation Ontario (2014) note that disobeying traffic signals accounts for 42% of fatal crashes and 29% of injury crashes.

Traffic enforcement is one way the authorities can mitigate the incidences and severity of red-light running and speeding. Thus, the subsequent sections shall discuss the implementation of *MPE* technologies in improving safety.

1.2 MPE Operation

Several cities have embraced the *MPE* technology and have registered positive results (i.e. limiting speeding and speed-related collisions). For example, the technology has reduced non-fatal and fatal crashes in France by 26 and 21%, respectively (*15*). Additionally, in Charlotte, North Carolina, the introduction of *MPE* has reduced collision cases by an average of 10%. In Washington, speeds were reduced by 14%, with a further reduction of 82% based on the number of vehicles exceeding the recommended 10 mph (*13*). In British Columbia, speed-related collision numbers shrank by 25% in enforcement locations (*15*). Moreover, there was a 22% reduction in collisions in Australia and a further 22% reduction in various collision cases, culminating in a 38% fall of collision-related injuries (*15*).

However, despite the reported success cases worldwide, it is unclear how the technology's efficacy in enhancing safety is recorded. Ordinarily, the primary concern has been allocating resources to obtain maximal safety impact (*15*). While comprehensive studies have focused on the methodologies, procedures, and performance measures applied in verifying the effectiveness of *MPE* programs, there are limited studies that underscore the systematic design processes that control the initialization and operations of the technology (*15*). Thus, this research aims to incorporate the program performance information in the literature alongside the data obtained from its enforcement to establish a robust study on the *MPE*'s effectiveness in the City of Edmonton.

1.3 Safe Mobility Strategy

In 2015, the City of Edmonton (*CoE*) adopted Vision Zero, aiming for zero traffic-related fatalities and severe injuries and based on the belief that traffic collisions are preventable and predictable, all road users make mistakes, and life loss is unacceptable. Thus far, fatalities have decreased by 56% and serious injuries by 30%. To achieve Vision Zero's goal and provide safe and livable streets, the *CoE* has launched the latest Safe Mobility Strategy 2021-2025. This strategy includes three central technical studies and papers: policy and planning, changing the conversation around traffic safety, and crash and equity analysis.

High injury locations are identified through the crash and equity analysis to help distinguish collision causes and identify vulnerable road users at increased risk of crashing. This step supports the introduction of proper countermeasures. This strategy considers all transportation modes, focusing on providing equal protection for all travellers, including walkers and cyclists unprotected by vehicle frames.

Achieving Vision Zero is not only the responsibility of the *CoE*'s Safe Mobility section but also an integrated work of other partners in the community. Therefore, the Safe Mobility Strategy requires the bearing of other groups such as Alberta Health Services (*AHS*), Edmonton Police Services (*EPS*), researchers in post-secondary institutions, and largely responsible citizens. The *CoE* seeks to achieve the Vision Zero target of zero fatalities and severe injuries by 2032.

1.4 Project Objective

To mitigate some of the existing speeding and safety concerns in the *CoE*, the main objective of this project is to correlate the deployed *MPE* hours and visits with the duration between two consequent collisions. The MPE and crash data are analyzed using survival analysis to achieve this goal. The project proposes the optimal *MPE* variable (i.e., hours, visits, or hours per visit) that provides the longest duration between collisions that correspondingly reduce the hazard of collision occurrence. Moreover, the analysis predicts the expected reduction of collision occurrence by deploying *MPE* variables. These project outcomes help the *CoE* better understand the optimal *MPE* deployment strategies for different sites in the city.

1.5 Report Structure

The remainder of this report is divided into five sections. The following section provides a literature review detailing the effectiveness of *MPE* and the use of survival analysis. The next section discusses the data used in this study. The subsequent section summarizes the project phases, including preparing the data and applying the survival analysis. The major results of the analysis follow with an in-depth discussion of its implications. The final section summarizes the conclusions and outlines potential future work.

2 LITERATURE REVIEW

This literature covers the effectiveness of *MPE* on collisions and speeding by underscoring the particular deterrence effects, deployment strategies, and resource allocation. Most of the studies reviewed in this research concern the *MPE* programs' influence on collisions and vehicle speeds. Previous studies have shown that the practical application of *MPE* can minimize mean vehicle speeds by 2% (*15*). Also, related studies have pointed out that *MPE* reduces severe collision cases causing injuries and fatalities.

The usefulness of the *MPE* program can be attributed to the immediacy, unavoidability, and punishment severity, which influence driver behaviour and attitude based on the specific and general deterrence mechanisms. In other words, potential violators will adhere to the outlined standards when they observe fellow violators being punished, called general deterrence. Further research has linked general deterrence to dangerous driving education, *MPE*, and awareness campaigns. On the other hand, specific deterrence is where a driver receives a firsthand experience, including detection and punishment (*17*). Li and Guo (2015) postulate that since general deterrence is typically compared to specific deterrence, the enforcement authorities should strive to meet the greater general deterrence (*17*). Some scholars posit that general deterrence can be enhanced by targeting high-risk periods and locations through non-visible and visible enforcement strategies to enhance unpredictability and enforcement publicity while also embracing long-term schemes in the enforcement program (*15*). In Edmonton, for instance, research established that an increased number of issued tickets and enforcement sites simultaneously reduced speed-related collision cases (*8*).

The same report indicated that a reduction in collisions was linked to the *MPE* program's reliability, notably when implemented with higher location coverage, increased issued tickets, and consistent checks (*17*). Often, various jurisdictions spur the guidance and regulations of these programs, such as Alberta's Automated Enforcement Guidelines (*18*).

2.1 MPE Locations

Typically, *MPE* units are located at sites known for speed limit violations, collisions, and public complaints regarding speeding (*19*). Additionally, they can be deployed when officers receive a special request from local government. Also, there are cases where the conventional speed control measures have failed or are infeasible (*19*).

Studies show that public awareness is effective for speed management. For instance, speed management programs in Alberta and Manitoba strive to enlighten the public on the problem and associated dangers linked to red-light running and speeding (*19*). Such programs aim to change the perceptions and attitudes of drivers by raising awareness of the risks of disregarding traffic rules (*15*). In addition, reports have shown how public awareness can be effective when complemented in synergy with other enforcement techniques (*15*).

Carnis (2011) highlights those debates over privacy, reliability of cameras, fairness, and satisfactory revenue that can emerge in jurisdictions when speed cameras are installed. Additionally, there is some skepticism by segments of the population that moderate speeding accentuates crash risks. Reports on this issue show that public awareness programs could be instrumental in exposing such beliefs (*20*) by targeting drivers through visible enforcement programs, awareness, and education efforts.

Reports looking into the effect of the yellow light phase on stopping behaviour reveal that this approach helps eradicate red-light violations. One report cites that an increase of one second in yellow duration, as long as it does not exceed 5.5 seconds, will decrease red-light violations by at least 50 % (*20*). However, the yellow light phase has few positive outcomes in reducing red-light crashes because drivers are more willing to try to get through the yellow light to pass the junction with a more extended light change interval (*20*).

Moreover, another study assessed *MPE* using the before-and-after Empirical Bayes (*21*). This method was applied on the urban arterial roads only, and it showed a reduction of 14% to 20% in collision severities by deploying speed enforcement. The researchers found the greatest effect of the *MPE* for severe collisions and, additionally, that continuous enforcement strategies are more functional than discontinuous ones.

2.2 MPE Program Framework

This section advances an *MPE* functional program design that entails four major critical stages for performance evaluation and resource allocation. The four stages include:

- Data gathering
- Site identification
- Enforcement scheduling
- Evaluation

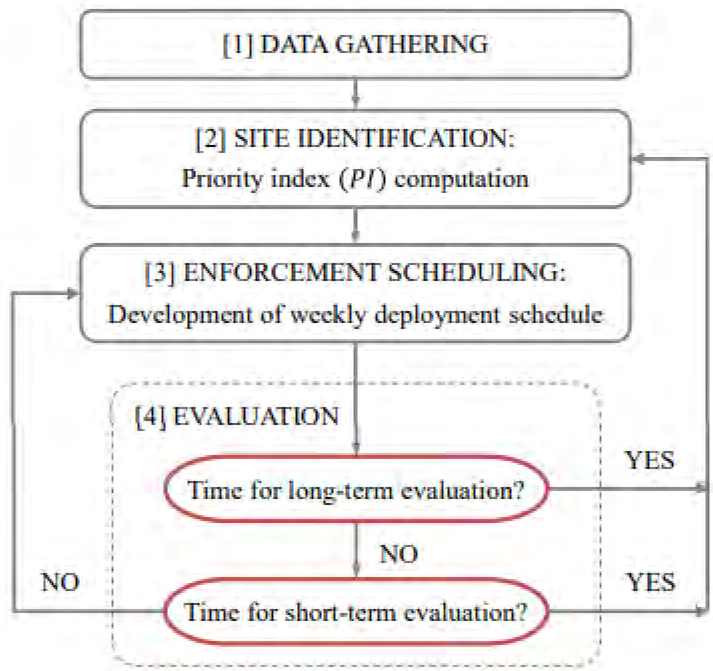


Figure 1*: MPE Program Framework in Edmonton, AB (Adapted from Kim et al., 2016).*

The previous framework shows how the *MPE* program is prepared and scheduled for different sites. First, it starts by collecting site data such as vehicle speed, speed limits, speed violations, and reported safety concerns. The next step is site identification, where the sites with higher cases of speed violations and collisions should prioritize deploying *MPE* (4). In other words, the frequency of crashes, road type, and related violations constitute the primary considerations when allocating the resources for deployment. Researchers should then evaluate the effectiveness of the deployed MPE program after either a short or long interval. The importance of this framework lies in determining the potential high-risk locations, thus, utilizing the *MPE* resources properly.

2.3 Survival Analysis & Hazard Models in Transportation Engineering

Survival analysis and hazard models have been widely used in medical applications while increasing for engineering applications. In general, survival analysis determines how long an event lasts, either in terms of "survival" or "failure," with failure often indicating structural failure in many engineering applications. However, in transportation engineering, survival analysis is commonly used in relation to the duration of roadway incidents, such as car crashes, impacting traffic (*22 - 26*).

There are three related functions in survival analysis: the failure function, the survival function, and the hazard function. The failure function illustrates the probability of a failure incident occurring before the specified point in time (*23*). Meanwhile, the survival function is the inverse of the failure function, illustrating the likelihood the duration continues beyond the specified point in time. The hazard function is related to, but distinct from, the other two functions; it represents the potential that an individual will “fail” at a specific point in time, having survived up until that point (*23*). For instance, Nam and Mannering examined the duration of traffic incidents from a dataset of incidents in Washington State during 1994-1995; these; these related to reporting, response, and clearance times by the Washington State Incident Response Teams. For their study, they considered the “failure” or the incident to be when an Incident Response Team cleared the incident and “survival” to be the persistence of the incident. The hazard function represented the probability that the incident would clear at any given time.

Previous studies show several variables can impact transportation engineering outcomes (i.e., incidents). For instance, both young and male drivers (i.e., gender and age variables) received speeding citations more often than other gender and age groups using Cox proportional hazard models (*27*). Moreover, drivers who receive speeding citations are at a higher risk of receiving more frequent speeding citations, which means speeding citations have little influence on changing speeding behaviour than other speeding penalties (*27*).

In addition, studies have used survival analysis to examine the vehicles’ mandatory lane-changing behaviours and related variables using the Cox proportional hazard model (*29*). Researchers showed that the type of vehicle has no significant impact on the duration of mandatory lane changing. Furthermore, the survival rate for mandatory lane-changing during the peak period is higher than the off-peak time (i.e., time is the considered factor).

A more detailed explanation of the Cox proportional hazard model can be found in (22, 24, 30, 31).

2.4 Summary

Speeding is considered a primary contributing factor to various types of collisions. Therefore, municipalities exert significant effort to deter speed violators using different tools and programs such as speed cameras, increasing the fines for speed violations, Mobile Photo Enforcement (*MPE*), and educational campaigns. *MPE* is an effective program to force drivers to follow traffic rules and drive according to the posted speed limit so as not to face charges and penalties. Previous studies explore transportation problems and the relevant factors that can help to improve road safety and the transportation system. For instance, different variables such as age, gender, time, and vehicle type have significantly affected speeding citations. Moreover, transportation engineering studies increasingly use survival analysis. By doing so here, this project aims to explore the effectiveness of *MPE* deployment variables in increasing the duration between collisions and consequently reducing the risk of being involved in a crash and, therefore, improving road safety.

3 DATA DESCRIPTION

The data used in the analysis was accessed through the Safe Mobility Section in the City of Edmonton (*CoE*) and provided in various Microsoft Excel files: "All-collision data," "*MPE* data," "Site," and "Event." The all-collision data spreadsheet included historical information about the collisions in Edmonton. These files combined data about collision cause, time, location, and travel direction. The *MPE* data file established the *MPE* control types and the start and the end of the *MPE* at each location. We extracted detailed information regarding the duration and number of *MPE* visits for each location from the spreadsheet. The control types used covered all the feasible methods of the *MPE* in the city. The third dataset was the Site spreadsheet, containing site IDs for all locations and a detailed description of each site. The "site ID" info is the core element to link all the files together, as will be clarified in the following sections. Finally, the Event dataset was a supporting file that provided an overall idea of the traffic count in each location.

The analysis period for this study is four years, from 2016 to 2019. Although data for the year 2020 was available, it was excluded from the study due to the documented impact of the *COVID-19* pandemic on the transportation system. The data were analyzed for each year separately and then integrated for further analysis. The following section will provide comprehensive details of the applied analysis method.

This work was carried out on various sites in Edmonton with different geometric characteristics and traffic volumes. All the available MPE control sites with sufficient amounts of the required information were analyzed based on the provided data. Therefore, the number of sites and locations differ from year to year based on the amount of information available.

4 METHODOLOGY

The procedural methodology for this study was achieved using four software programs: Microsoft Excel, MATLAB, SAS, and Stata. We executed the methods in two major stages: i) preparing the data and ii) applying the survival analysis. All the steps were examined for each year separately, followed by all years combined. We applied this process to explore the relationship between the frequency of deployed *MPE* (i.e., number of visits and hours) and the duration between two consequent collisions. Figure 2 shows the workflow summarizing the paper framework. The following sections explain the applied method in more detail.

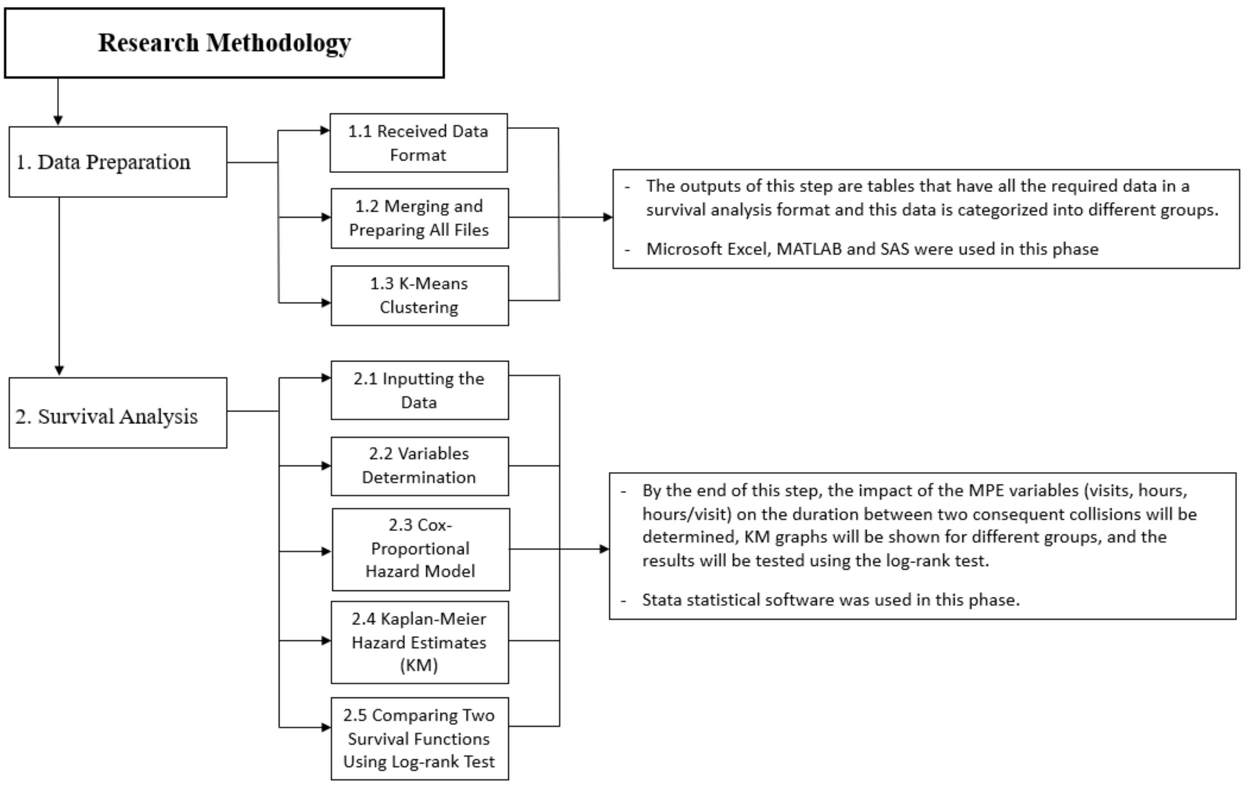


Figure 2: *Flowchart of the Methodology Employed*

4.1 Data Preparation

The Excel sheets requiring tabulation first were required to be used readily in the survival analysis (i.e., the second stage). This step aimed to merge all the data files into one, then change all the combined files into a survival analysis format. The output from this process was an Excel sheet that contained the essential info for each location (i.e., Site ID). While this step could be done manually, it would take excessive time to combine all files properly. Therefore, we scripted a

MATLAB code to speed up the processing time. The following subsections clarify this step thoroughly.

4.1.1 Received Data Format

Multiple Excel sheets contained a large set of data. The following bullets show the files and their contents.

- All collision data: this file included collision code, attachment code, collision key, collision cause, collision classification, collision data, collision location name, collision month, collision time, collision report year, travel direction, on-street name, and at the street name. Table 1 shows a sample of 'all collision data.'

***Table 1**: Sample of All Collision Data.*

#Collisions	COLLISION_CAUSE_NAME	COLLISION_CLASSIFICATION	COLLISION_LOCATION_NAME	COLLISION_MONTH	COLLISION_REPORT_YEAR	COLLISION_TIME	DAY_OF_MONTH	TRAVEL_DIRECTION
423332	FLWD TOO CLOSELY	PDO	JASPER AVENUE, 590 & 106 STREET NW	1	2017	830	6	NORTH
423333	FLWD TOO CLOSELY	PDO	JASPER AVENUE, 590 & 106 STREET NW	1	2017	830	6	NORTH
423334	STRUCK PRKD VEH.	PDO	107 AVENUE NW & 181 STREET NW	1	2017	1530	5	UNKNOWN
423335	STRUCK PRKD VEH.	PDO	107 AVENUE NW & 181 STREET NW	1	2017	1530	5	Not Applicable
423336	RAN OFF ROAD	PDO	90 AVENUE NW & 75 STREET NW	1	2017	1300	6	WEST

- Control – MPE: This file is essential as it provides data of control ID, site ID, violation category, control type, start date, end date, posted speed, and speed threshold. The data was filtered to contain only the *MPE* control type. Table 2 presents part of the *MPE* data.

Table 2: *Sample of MPE Control Data.*

Control Id	Site Id	Violation Category	Control Type	Start Date	End Date	Posted Speed	Speed Threshold
210934	211	Jenoptik	Jenoptik	17-Mar-17	18-Mar-17	100	115
212995	211	Jenoptik	Jenoptik	02-Apr-17	03-Apr-17	100	115
214060	211	Jenoptik	Jenoptik	10-Apr-17	10-Apr-17	100	115
214200	211	Jenoptik	Jenoptik	11-Apr-17	11-Apr-17	100	115
214924	211	Jenoptik	Jenoptik	17-Apr-17	17-Apr-17	100	115

- Site: site ID, location description, police division, direction, speed, double fine site, speed time site, speed posted, photo enforcement posted, and site type. Table 3 represents a data sample from the Site.

Table 3: *Sample of Site Data.*

Site Id	Location Description	Direction	Speed	Speed Posted	Photo Enforcement Posted	Site Type
101	100 Ave Between 160 - 162 St.	WB	50	1	1	Photo Radar Camera
102	Mark Messier Tr north of 137 Ave.	NB	70	1	1	Photo Radar Camera
103	130 St. between 111 - 113 Ave	NB	50	0	0	Photo Radar Camera
104	142 St at 95 Ave	NB	60	1	1	Photo Radar Camera
105	36 St between 119 - 121 Ave	NB	50	0	0	Photo Radar Camera

- Jenoptik Event: deployment, site ID, watch date, and traffic. This data was integrated to get the traffic count for each site. Table 4 is a sample of the data from the Jenoptik Event.

Table 4: *Sample of Jenoptik Event Data.*

Deployment	Site	Watch Date	Traffic
1965	2747	28-04-17	201
2870	2504	07-05-19	633
944	5445	05-10-17	328
1549	2503	04-02-17	1
5549	2552	20-12-19	288

4.1.2 Merging & Preparing All the Files

We scripted and tested a MATLAB code for the raw data files to overcome the difficulty of merging files manually. The code began by reading the input data files and storing them in four tables (All-collision data, Site IDs, *MPE* data, and Traffic data). Then it set up the Excel files to store the output and all the headers for the data. Following on, the code looped through all the Site IDs to determine locations for each, i.e. the street and avenue number/name along with the direction of travel (e.g. Site ID 104 represents the location of 142nd street and 95th avenue, Northbound direction). Thereafter, we applied a filter to extract all the collisions for particular locations within the given site description (i.e., street and avenue name and travel direction) and information to determine collision dates, times, and seasons (i.e., winter, summer, etc.) by looking at each collision individually and checking its location.

Using another loop, we generated a cycle through all the collisions to sort and calculate the number of hours visited, number of visits, traffic volume, and the time difference between collisions. We calculated the number of hours visited by filtering all the data from the *MPE* data table for those within the given year. From this, we could extract the time from the start to the end of the visit. We then calculated the number of visits by counting the total visits from the *MPE* data table that fell within the time range. The traffic count was calculated by filtering all the data from the traffic data table to search only for the ones within the given year and extract the amount of traffic.

All these loops were executed on all the site IDs for each year. Lastly, we stored all the calculated data in an Excel spreadsheet by year (i.e., extracted an Excel sheet for each year). The average processing time for this step was 90 minutes. The code outputs were visually inspected to verify accuracy and matched the expected results. Table 5 shows a sample of the output of this step for 2019.

Table 5*: Sample of the Output Data from The Data Preparing Phase.*

Site ID	Days b/w Coll.	Event Occ.	1st Coll. Date	2nd Coll. Date	# Times visited	# Hours	Hours/ visit	1st coll. Season	2nd coll. Season	Duration Season	Sev.	Coll. Cause
2449	233.7	1	01-Jan-2019	22-Aug-2019 16:59:59	144	453.43	3.15	Winter	Summer	Transition	PDO	FLWD TOO CLOSELY
2449	38	1	22-Aug-2019 16:59:59	29-Sep-2019 16:59:59	1	3.62	3.62	Summer	Winter	Transition	Min.	FLWD TOO CLOSELY
2449	92.3	0	29-Sep-2019 16:59:59	31-Dec-2019	42	124.77	2.97	Winter	Winter	Winter	N/A	N/A
2452	53.7	1	01-Jan-2019	23-Feb-2019 16:59:59	5	13.50	2.70	Winter	Winter	Winter	PDO	LFT TURN X PATH
2452	310.3	0	23-Feb-2019 16:59:59	31-Dec-2019	56	144.68	2.58	Winter	Winter	Winter	N/A	N/A

4.1.3 *K*-Means Clustering

We used *K*-means clustering to categorize the sites based on the *MPE* number of hours, *MPE* number of visits, *MPE* hours/visit, and traffic count. This form of clustering groups observations together with similar characteristics (*32*), i.e. into clusters in which each observation belongs to the cluster with the nearest mean. First, the number of clusters (*K*) should be determined; in our work, we used two clusters to categorize the variables into two groups: above-average and below-average. Then, two cluster seeds are randomly specified as the clusters' centroids. After that, each observation is assigned to one of the clusters based on its proximity that had the least squared Euclidean distance. Finally, the centroid of each cluster is calculated, and iterations are done until convergence is reached (i.e., the same points are assigned to the same cluster in repetitive cycles). We used SAS software to apply the *K*-means clustering process. Though the *K*-means was conducted in SAS, we entered and clustered the data into two groups. Figure 3 summarizes the *K*-means clustering algorithm, and Table 6 shows the clustered groups.

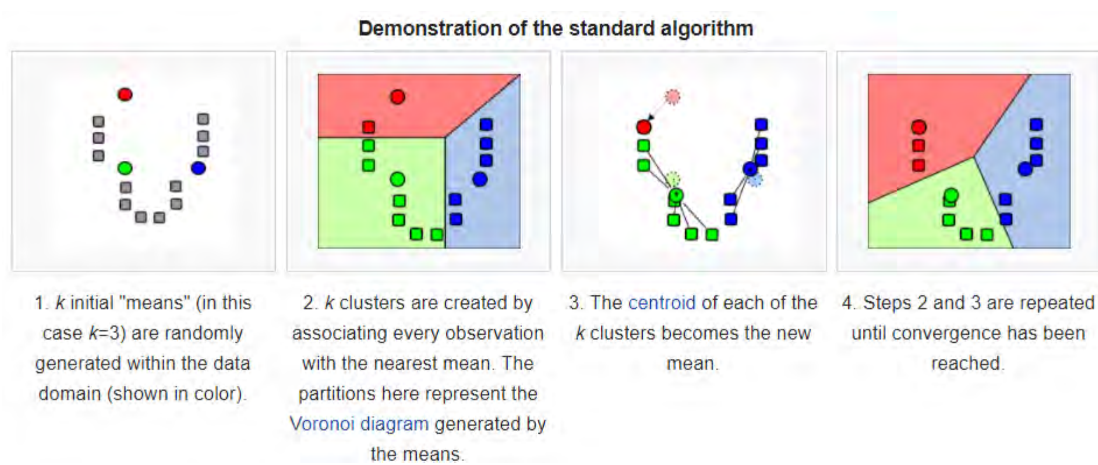


Figure 3: *Demonstration of the Standard Algorithm (source: Wikipedia).*

***Table 6**: The Clustered Groups.*

Variable	Group 1	Group 2
#*MPE* hours	Above-average *MPE* hours	Below-average *MPE* hours
#*MPE* visits	Above-average *MPE* visits	Below-average *MPE* visits
#*MPE* (hours/visit)	Below-average *MPE HpV*	Above-average *MPE HpV*
Traffic volume	Below-average traffic volume	Above-average traffic volume

4.2 Survival Analysis Method

The second phase in the methodology was to apply the survival analysis process. We chose survival analysis to study the time to event occurrence. Thus, in this case, the failure event was the collision occurrence. The basic concepts of the survival analysis are to define the hazard and survival functions, create the Kalan-Meier survival curves for different variables and compare two survival curves using the log-rank test. We used Stata Statistical Software to execute this step. Our survival analysis entailed five stages. By the end of this process, the relationship between the *MPE* variables (number of hours, number of visits, and the ratio hours/visit) and the duration between two consequent collisions could be determined. The following subsections will illustrate the implemented phases.

4.2.1 Inputting the Data

The data must first be classified as survival-time data and tabulated to include the time and failure variables to begin this process. The time variable represents the duration between two consequent collisions in days, and the failure variable is a binary value, i.e., 0 or 1. The failure value is 0 when there is no collision and 1 when there is a collision during the specified period. Moreover, the data in this step (Table 4) includes site ID, duration between every two consequent collisions, first and the second collision dates used to estimate the time between collisions, number of deployed *MPE* visits, number of deployed *MPE* hours, the ratio between the number of *MPE* hours to the number of *MPE* visits (*HpV*), first and second collisions occurrence season, collision severity, collision cause, land-use, and traffic count.

4.2.2 Variables

The variables employed here were the number of deployed *MPE* visits, number of deployed *MPE* hours, the ratio between them (*HpV*), and traffic count. We will examine the impact of these variables on the duration between two consequent events (i.e., collisions) in the survival analysis phase.

4.2.3 Cox-proportional Hazard Model

The cox-proportional hazard model investigates each variable's effect on the time between collisions. This model is the most popular technique of the semi-parametric methods since it does not make a hazard baseline assumption, a benefit when choosing a predictive model (*33*). It explores the rate of a specific event occurrence (i.e., hazard rate) as an influence on different factors.

The general Cox proportional hazard model is;

$$h\ (t, x) = h_0\ (t) \exp(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_i x_t)$$

The Cox regression prototype can be changed into another equation by the logarithmic transformation (Shi et al., 2014, p. 33);

$$ln\ ln\ \frac{h\ (t,x)}{h_0\ (t)} = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_i x_t$$

The relative risk *RR* can be denoted as $\frac{h\ (t,x)}{h_0\ (t)}$, then, the *COX* regression is the linear model of the logarithm of the *RR*. Under other covariate variables remaining constant, βi shows the logarithm changes of the *RR* with the unit change of the i^{th} covariate variable. Based on the definition above, the *COX* regression has the following properties: 1) If βi>0, the i^{th} variable is a risk factor, and its hazard may be higher with the increasing time. And this indicates that the incident may be disposed of quickly. 2) If β<0, it means this variable is a protective factor, and the duration of the traffic incident is longer, which indicates the incident cannot be disposed of in time. 3) If β=0, this variable has nothing to do with the traffic incident duration.

The outcome of this step is a regression model that relates the deployed *MPE* hours, visits, and *HpV* separately. These models provide the hazard ratio for each variable; this report will now explore the impact of the variables above.

4.2.4 Kaplan-Meier Hazard Estimates (KM)

The *KM* hazard estimate is a univariate nonparametric analysis used to estimate the survival probability from observed survival times (*34*). Our study established the *KM* survival curve to differentiate the impact of MPE variables on different road categories. It is generated by taking the product of conditional probabilities sequences and obtaining the standard estimator (i.e., *KM* estimator). The complete details of how the calculation is done have been provided by (*32)*. As shown in the following figure, this curve consists of a series of steps, and each step represents an event occurrence (i.e., collision). The survival probability is on the *Y*-axis of the curve, and the time duration is on the X-axis. Thus, the cumulative survival probability can be extracted at any time point by obtaining the corresponding value on the *Y*-axis. The estimated cumulative survival at any time point is at 95% confidence intervals.

If T_i represents the event period of the traffic of the ith term, and the time-series of the traffic event in the state is $T_1 < T_2 <\ldots T_n$, then the Kaplan-Meier based survival likelihood of traffic occurrence period represented by Š (t) is given by;

$$\check{S}(t) = \prod_{T_i^c \Sigma \quad t} \frac{n-i}{n-i+1}$$

In this case, the traffic event time is represented by T_i^c of the ith term entire samples. However, for a sample to be complete, it has to attain certain conditions, including, T_i^c is less than t besides being a positive integer, where $T_i^c \leq t$ and$T_i^c \in Z$ (*36*).

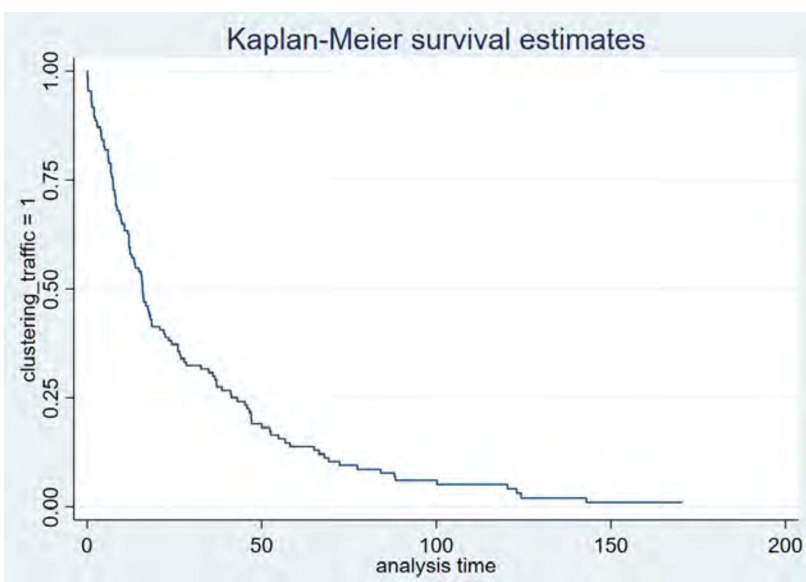


Figure 4*: Example of KM Graph.*

The curve shown in Figure 4 provides a beneficial data summary that can be utilized to estimate measures as median survival time. It plots the difference between the survival probability for two comparable groups (i.e., clustered variables). The data of these groups should be in a categorical format as this method estimates the cumulative survival probability for each group separately. After being clustered, we determined these groups based on the previously outlined *K*-means Clustering method performed in Stage 1.

Therefore, this step's output is a graph for each variable that shows the survival probability at any time point during the analysis period. Also, it can be used to compare the survival probabilities for different categorical groups. The survival probability for each group is generally checked against the mean value (50%) to facilitate comparison between the groups.

4.2.5 Comparing Two Survival Functions Using Log-rank Test

There are two methods to compare survival functions; the first method uses a prespecified time point; the other compares the overall survival experience, called the log-rank test. The log-rank test is considered more reliable than the prespecified time point method for reasons outlined in (*37*). As a result, our study applied and executed the log-rank test across the entire survival time range. The test null hypothesis between the two groups is:

$$H0\colon\ S1\,(.)\ =\ S2\,(.)$$

where the dot represents the whole survival time range.

The alternative hypothesis is applicable when this null hypothesis is rejected. Moreover, the log-rank test compares the observed and the expected number of collisions if the two groups have the same survival function. Thus, if the null hypothesis is true, the two groups would have the same survival probability, determined based on the *p*-value. The Chi-square value is compared for each group using the standard Chi-square test. By the end of this step, the equity test of the two groups shall determine whether or not they have the same survival time probability. In this study, the log-rank test is used as a validator for the previous steps as it shows whether the different road categories have the same survivability or not.

5 DATA ANALYSIS & RESULTS

The methodology involved two major stages, as shown in Figure 2. The following section discusses the results of the implemented procedure, shown for each year separately. In addition, the analysis for each year is executed on different groups, as outlined in the following subsections.

5.1 Survival Analysis Results (2019)

5.1.1 All Sites

We applied the methodology to all sites from Jan 1, 2019, to Dec 31, 2019. There were 250 sites in Edmonton that had an *MPE* during this period. These sites experienced 573 collisions, mainly classifiable as Property Damage Only (*PDO*). In addition, there were seven major collisions and fifty-two minor collisions. The primary cause was following too closely, often coupled with speeding, considered a speeding-related collision cause. For deployed *MPE* between two collisions, the number of deployed *MPE* visits ranged between 0 to 195, the number of deployed *MPE* hours varied between 0 to 620, and the ratios between hours and visits were estimated as low as 0 and high as 4 hours/visit.

For the all-sites group, we conducted the Cox *PH* models. We then plotted the Kaplan-Meir graphs for different variable groups. Finally, we undertook the log-rank tests. Table 7 a-c summarises the results:

5.1.1.1 *Cox-proportional hazard model*

Table 7*: The Hazard Ratio Estimates for the MPE Variables (All Sites).*

	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Hours	0.9942	0.0008	0.000	0.9925	0.995

(a) The total number of deployed *MPE* hours.

	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Visits	0.9795	0.0026	0.000	0.9742	0.9848

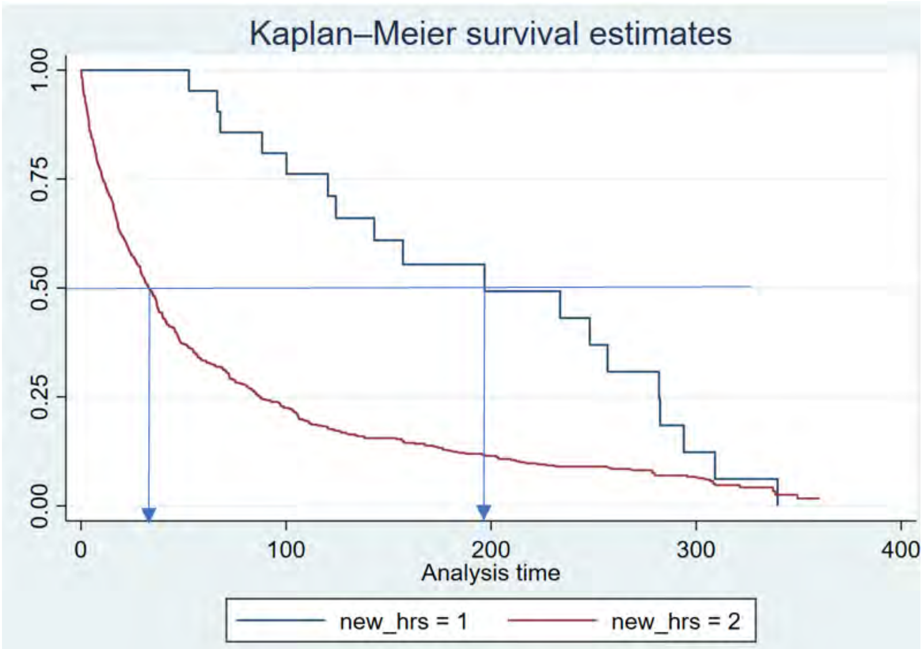
(b) The total number of deployed *MPE* visits.

	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Hours per Visits	0.7861	0.0235	0.000	0.7413	0.8336

(c) The ratio between the number of *MPE* visits and hours (*HpV*).

The results of this group show that there is significant evidence of the *MPE* impact on increasing the duration between two consequent collisions. The deployed *MPE* hours and visits showed only small percentage reductions (1% and 3%, respectively) in the collision hazards. Given the percentage difference between the two, one conclusion might be that investing in increased visits would produce better outcomes than increasing the number of hours. For instance, if the *CoE* invested four *MPE* hours, the more significant benefit would result from splitting these into different shorter visits. In addition, the ratio of deployed *MPE* hours to visits (*HpV*) has a hazard ratio (*HR*) of 0.78, with a reduction of 22% in collision occurrences for locations that had a high *HpV* compared to sites without *MPE*.

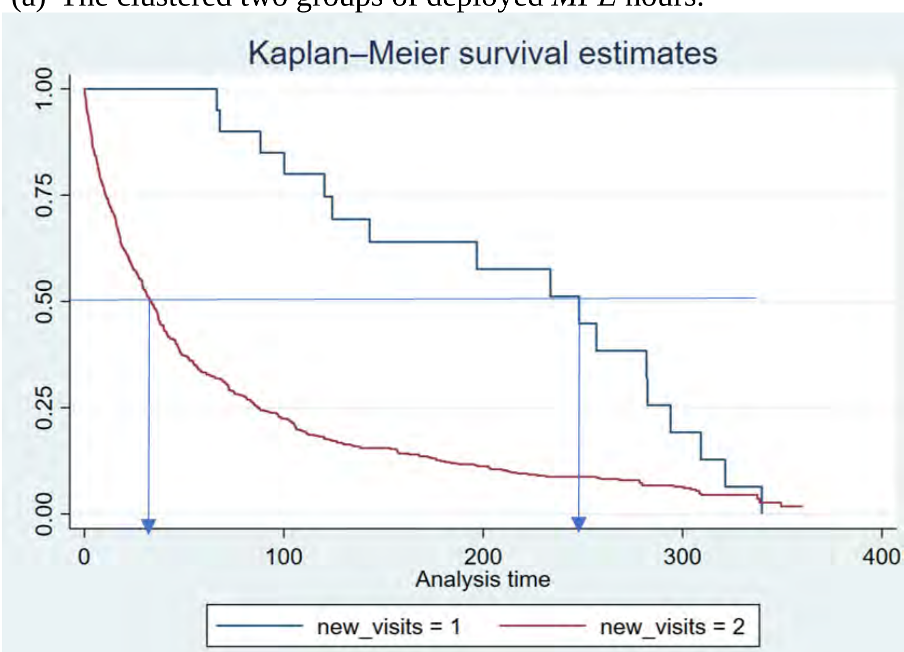
5.1.1.2 Kaplan-Meier Graphs



1 = Above-average

2 = Below-average

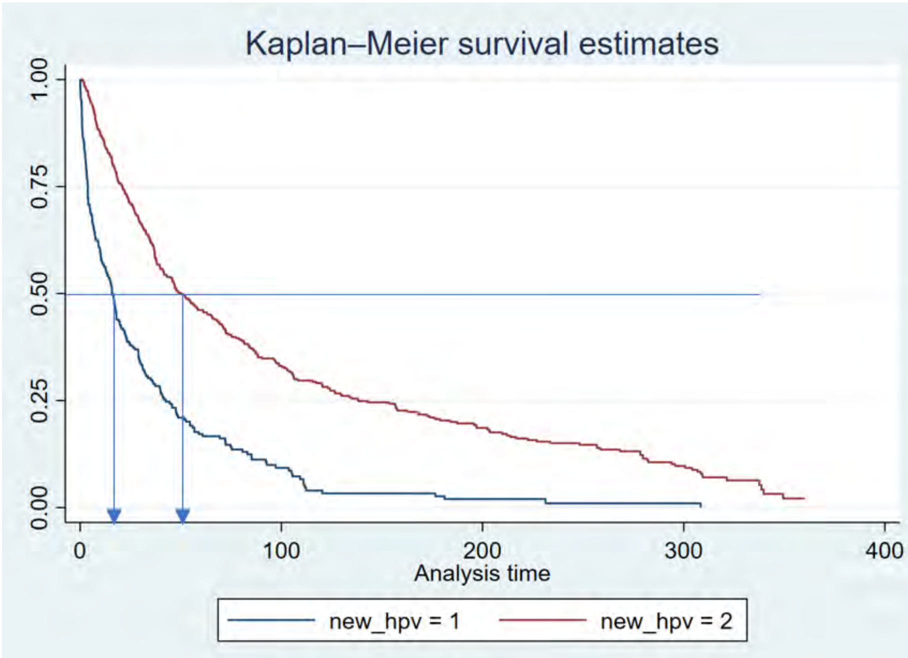
(a) The clustered two groups of deployed *MPE* hours.



1 = Above-average

2 = Below-average

(b) The clustered two groups of deployed *MPE* visits.

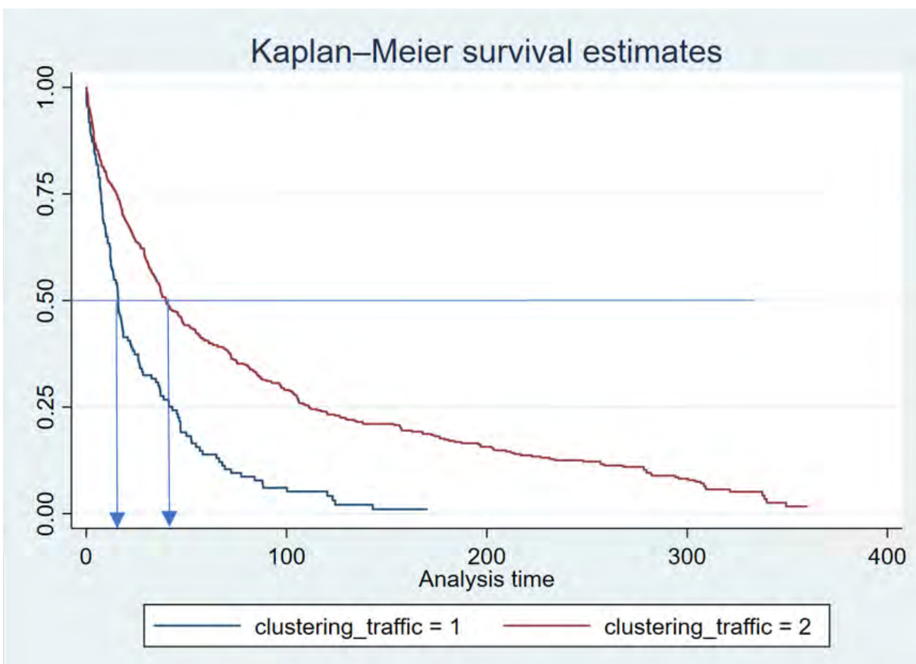


1 = Below-average

2 = Above-average

(c) The clustered two groups of deployed *MPE* hours/visit.

Figure 5: *The KM Survival Estimates for The MPE Groups (All Sites).*



1 = Below-average

2 = Above-average

Figure 6*: The KM Survival Estimates for the Clustered Traffic Groups (All Sites).*

We compared the survival probability for different locations that experienced two levels of deployed *MPE* (i.e., above-average deployed *MPE* and below-average deployed *MPE*) by plotting the *KM* graphs. First, the KM graphs yield the same results regarding the deployed MPE hours and visits, showing that the locations that experienced a high level of *MPE* hours or visits (above average) have higher survivability than the locations with lower *MPE* hours or visits (below average). As shown in Figure 5a, the median survival probability (at 0.5 on Y-axis) for sites with above-average deployed *MPE* hours is 198 days. In comparison, the below-average ones had a survival probability of 38 days. Similarly, in Figure 5b, the survival probability at 0.5 on Y-axis for sites that had above-average *MPE* visits is 248 days; meanwhile, the below-average group is 37 days. Moreover, the groups with below-average *MPE* hours and visits encountered more frequent steps in the *KM* graphs, indicating that these sites experienced more collisions over a short period than the above-average *MPE* groups.

Second, Figure 5c shows that, for the clustered *MPE* hours/visit groups, the sites that experienced a higher ratio of the deployed *MPE* hours to visits have more survival probability than the below-average locations. Finally, Figure 6 illustrates that above-average traffic volume locations have higher survivability than the below-average sites since drivers tend to speed up when roadways have little or no congestion.

5.1.1.3 Log-rank Equality Test

***Table 8**: The Log-Rank Test for Different Groups (All Sites).*

Hours	Observed Events	Expected Events
1 2	18 555	41.37 531.63
Total	573	573.00
		Chi2(1) = 14.61 Pr>chi2 = 0.0001

a) Log-rank test for *MPE* hours groups

Visits	Observed Events	Expected Events
1 2	17 556	41.74 531.26
Total	573	573.00
		Chi2(1) = 14.34 Pr>chi2 = 0.0001

b) Log-rank test for *MPE* visits groups

HpV	Observed Events	Expected Events
1 2	228 345	126.12 446.88
Total	573	573.00
		Chi2(1) = 111.63 Pr>chi2 = 0.0000

c) Log-rank test for *MPE* hours/visit groups

Traffic Count	Observed Events	Expected Events
1 2	123 450	66.91 506.09
Total	573	573.00
		Chi2(1) = 55.29 Pr>chi2 = 0.0000

d) Log-rank test for traffic volume groups

We conducted the log-rank tests to see if two different groups have the same survivability (i.e., null hypothesis) and whether those comply with the results from the *KM* graphs. Thus, if the *Pr* value is less than 0.05, the tested groups do not have the same survival probability. As shown in Table 8, all of the tested groups had $Pr < 0.05$, which proves the previous results of the *KM* graphs.

5.1.2 Arterial & Collector Roads

In this subsection, the analysis process is executed on the arterial and collector sites only. We tested this category to examine the impact of *MPE* on the main roadway categories in Edmonton. There were 111 arterial and collector sites with data available for 2019. These locations had 566 collisions, with the two leading causes identified as following too closely and left turn crossing path. These collisions resulted in seven major and fifty-two minor crashes. Table 9 and Figure 7 a-c show the analysis results, followed by the discussion of the results.

5.1.2.1 Cox proportional hazard model

***Table 9**: The Hazard Ratio Estimates for the MPE Variables (Arterial and Collector Locations Only).*

	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Hours	0.99414	0.00085	0.000	0.99247	0.9958

(a) The total number of deployed *MPE* hours.

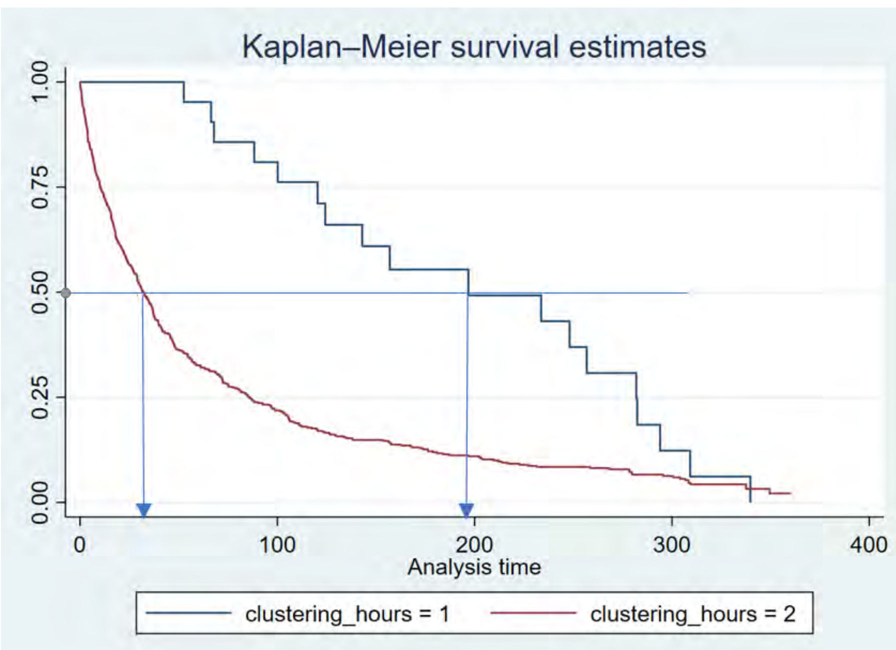
	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Visits	0.9791	0.00275	0.000	0.9737	0.9845

(b) The total number of deployed *MPE* visits.

	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Hours per Visits	0.78519	0.0234	0.000	0.7405	0.8324

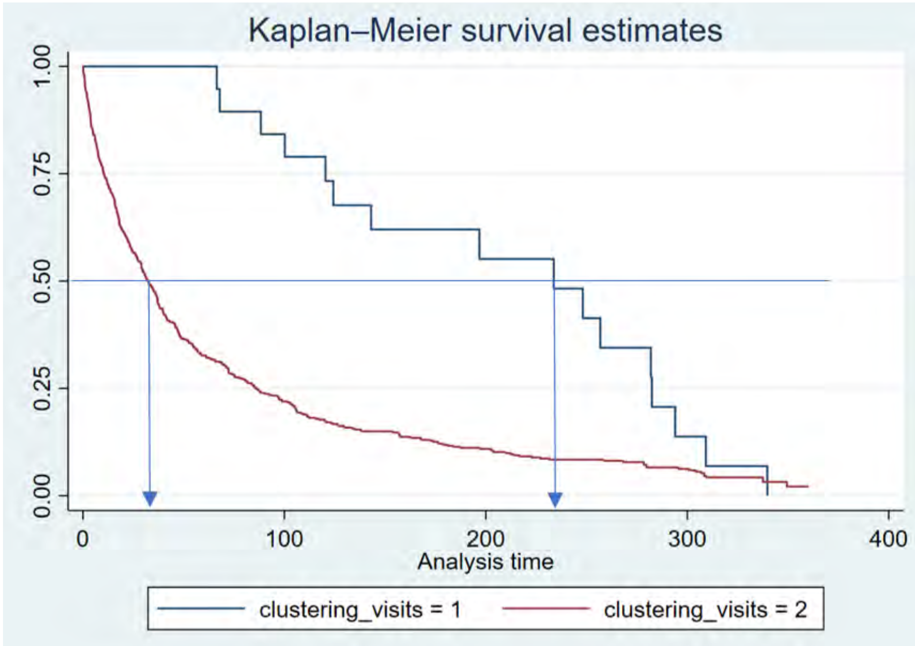
(c) The ratio between the number *MPE* of visits and hours (*HpV*).

5.1.2.2 Kaplan-Meier Graphs



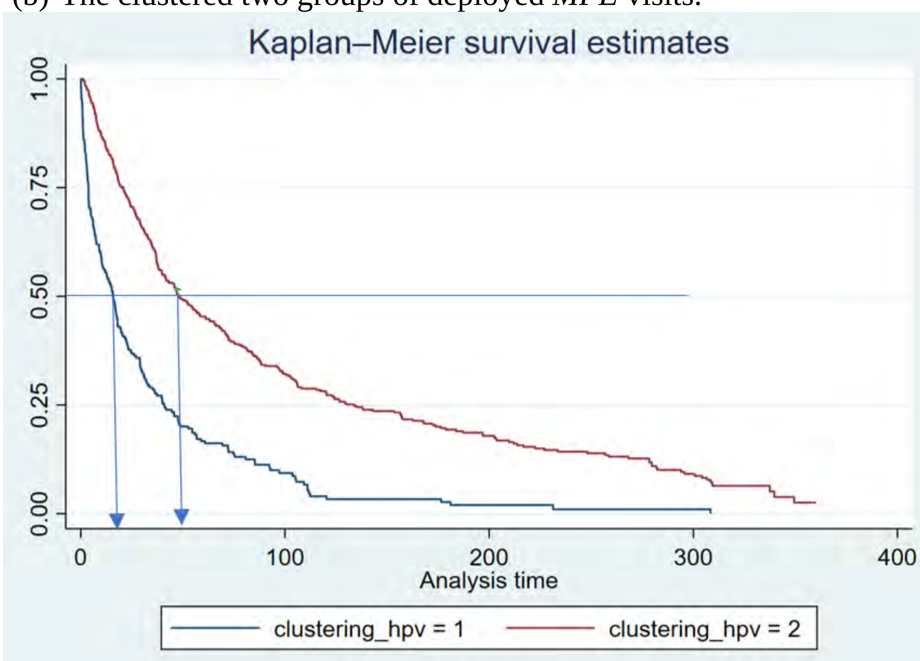
1 = Above-average
2 = Below-average

(a) The clustered two groups of deployed *MPE* hours.



1 = Above-average
2 = Below-average

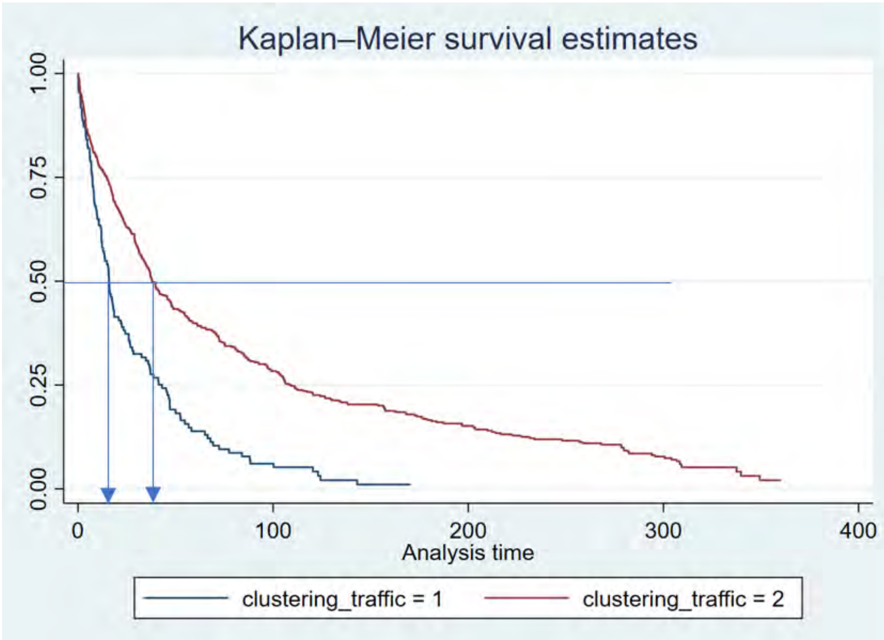
(b) The clustered two groups of deployed *MPE* visits.



1 = Below-average
2 = Above-average

(c) The clustered two groups of deployed *MPE* hours/visit.

Figure 7: *The KM Survival Estimates for The MPE Variables' Groups (Arterial and Collectors Locations Only).*



1 = Below-average
2 = Above-average

Figure 8: *The KM Survival Estimates for The Clustered Traffic Groups (Arterial and Collector Locations Only).*

5.1.2.3 Log-rank Equality Test

***Table 10**: The Log-Rank Test for Different Groups (Arterial and Collectors Only).*

Hours	Observed Events	Expected Events
1 2	18 548	41.97 524.03
Total	566	566.00
		Chi2(1) = 15.21 Pr>chi2 = 0.0001

a) Log-rank test for *MPE* hours groups.

Visits	Observed Events	Expected Events
1 2	16 550	39.15 526.85
Total	566	566.00
		Chi2(1) = 15.14 Pr>chi2 = 0.0001

b) Log-rank test for *MPE* visits groups.

HpV	Observed Events	Expected Events
1 2	225 341	124.59 441.41
Total	566	566.00
		Chi2(1) = 109.52 Pr>chi2 = 0.0000

c) Log-rank test for *MPE* hours/visit groups.

Traffic Count	Observed Events	Expected Events
1 2	123 443	68.26 497.74
Total	566	566.00
		Chi2(1) = 51.82 Pr>chi2 = 0.0000

d) Log-rank test for traffic volume groups.

The outcomes of the *Cox PH* models provide almost the same results for all sites. The results for all the *MPE* variables are significant at a 95% confidence interval and had *HR* values less than 1, which reflects the positive impact of deployed *MPE* variables. The *HR* for the deployed *MPE* hours/visit surpassed other *MPE* variables as it resulted in a reduction in collisions occurrence of 22%.

The *KM* survival estimates for the arterial and collector locations yielded the same conclusion for all the sites' results and followed similar trends. For instance, in Figure 7, sites that experienced above-average *MPE* hours, visits, or *HpV* possess a higher survival probability than the below-average category. In addition, the below-average traffic volume locations have a higher risk of collision occurrence, as shown in Figure 8.

The log-rank tests examined the equity of the survival probability for the clustered groups. The *Pr* values for all tests were significant (i.e., $Pr < 0.05$). Thus, the log-rank test results compare favourably with the outcomes of *KM* graphs.

5.1.3 Arterial Sites Only

We tested the arterial sites separately to examine the impact of *MPE* variables. Sixty-four arterial sites collectively recorded 436 collisions in 2019. These collisions' severity was mainly *PDO*, with four major and thirty-eight minor collisions. The leading causes for the arterial road collisions were following too closely, left turn cross-path, and improper lane change. Furthermore, the deployed *MPE* hours and visits between collisions ranged from 0 to 620 and 0 to 195, respectively. Table 11 a-c, Figure 9 a-c, Figure 10, and Table 12 a-d illustrate the analysis results.

5.1.3.1 Cox proportional hazard model

Table 11*: The Hazard Ratio Estimates for the MPE Variables (Arterials Only).*

	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Hours	0.9939	0.00085	0.000	0.99230	0.99567

(a) The total number of deployed *MPE* hours.

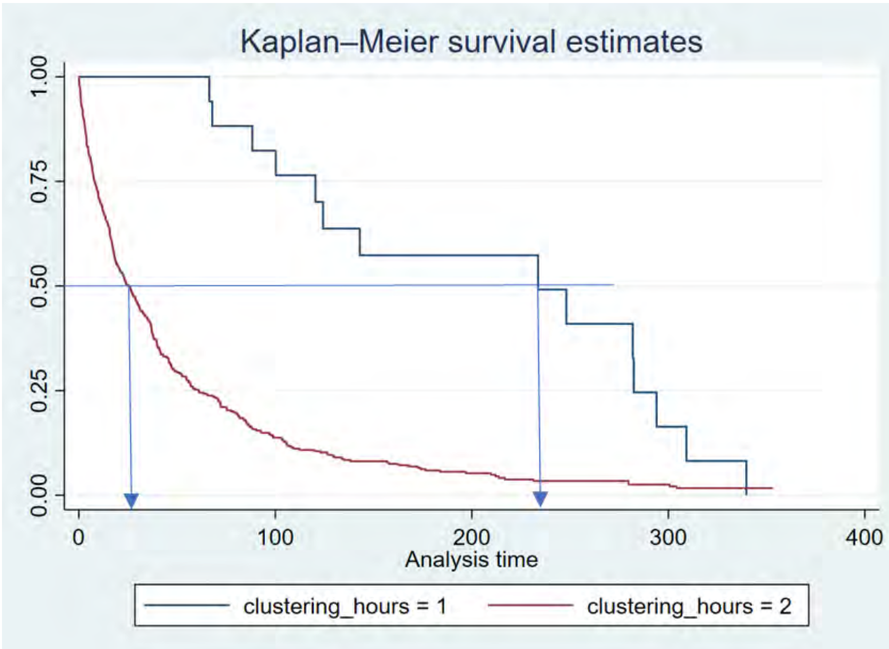
	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Visits	0.9798	0.00278	0.000	0.9743	0.9852

(b) The total number of deployed *MPE* visits.

	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Hours per Visits	0.78457	0.0255	0.000	0.73609	0.8362

(c) The ratio between the number of *MPE* visits and hours (*HpV*).

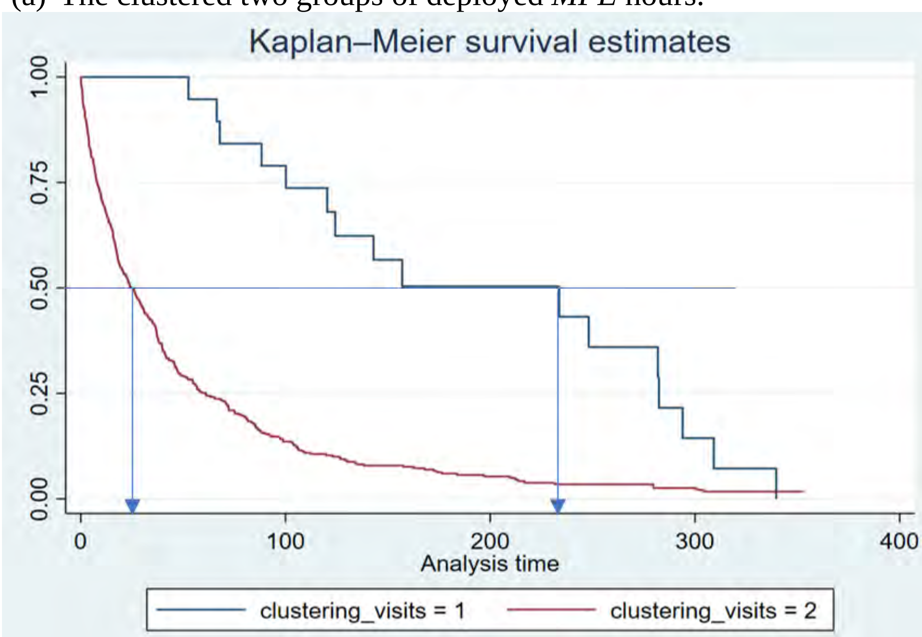
5.1.3.2 Kaplan-Meier Graphs



1 = Above-average

2 = Below-average

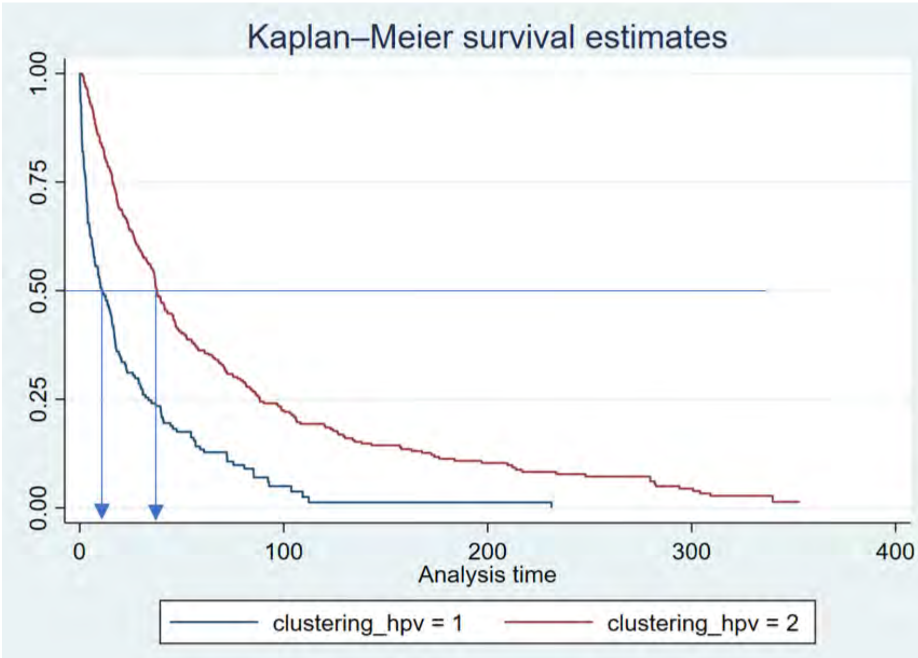
(a) The clustered two groups of deployed *MPE* hours.



1 = Above-average

2 = Below-average

(b) The clustered two groups of deployed *MPE* visits.

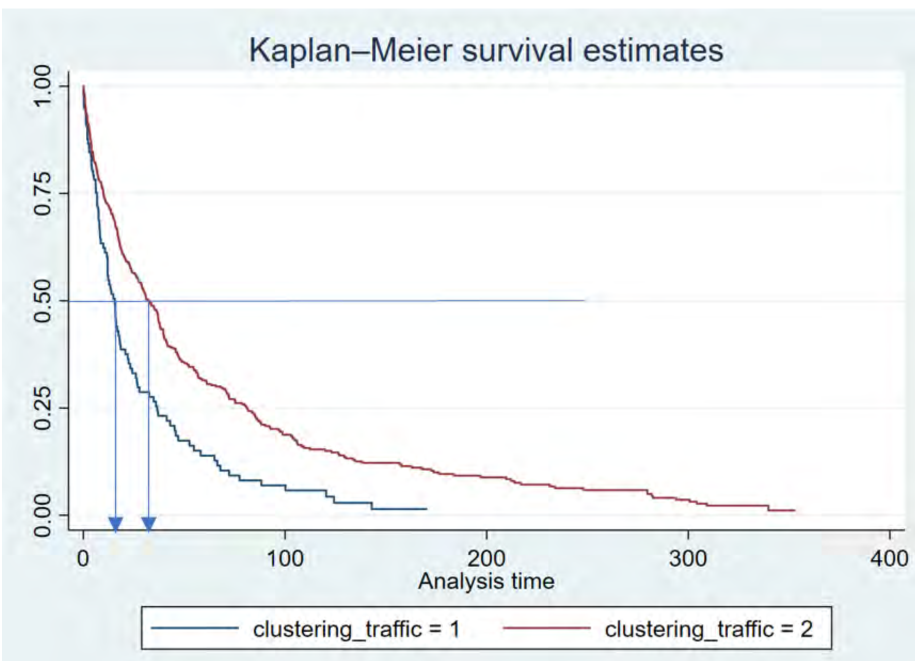


1 = Below-average

2 = Above-average

(c) The clustered two groups of deployed *MPE HpV*.

Figure 9: *The KM Survival Estimates for the MPE Groups (Arterials Only).*



1 = Below-average

2 = Above-average

Figure 10: *The KM Survival Estimates for the Clustered Traffic Groups (Arterials Only).*

5.1.3.3 Log-rank Equality Test

***Table 12**: The Log-rank Test for Different Groups (Arterials Only).*

Hours	Observed Events	Expected Events
1 2	14 422	43.37 392.63
Total	436	436.00
		Chi2(1) = 23.98 Pr>chi2 = 0.0000

a) Log-rank test for *MPE* hours groups.

Visits	Observed Events	Expected Events
1 2	16 420	46.85 389.15
Total	436	436.00
		Chi2(1) = 24.63 Pr>chi2 = 0.0000

b) Log-rank test for *MPE* visits groups.

HpV	Observed Events	Expected Events
1 2	164 272	92.37 343.63
Total	436	436.00
		Chi2(1) = 74.16 Pr>chi2 = 0.0000

c) Log-rank test for *MPE* hours/visit groups.

Traffic Count	Observed Events	Expected Events
1 2	90 346	57.40 378.60
Total	436	436.00
		Chi2(1) = 21.85 Pr>chi2 = 0.0000

d) Log-rank test for traffic volume groups.

The Cox PH models showed that the MPE variables had a proactive effect on collision occurrence compared to locations not exposed to *MPE* deployment. For example, the *deployed MPE hours/visit HR* was 0.78, which means an expected reduction in the risk of collision occurrence of 22%. These results are significant at a 95% confidence interval.

The *KM* survival graphs, shown in Figure 9, emphasize the overall conclusion that the higher the level of deployed *MPE*, the less risk of collisions occurrence. To illustrate, locations that had above-average *MPE* hours had an average survival probability of 231 days, while below-average *MPE* hours sites survived for an average of 27 days. In other words, the below-average *MPE* locations experienced more frequent collisions. Moreover, similar to all sites, the survival probability for above-average traffic volume locations was higher than those below-average, possibly due to changes in drivers’ speeding behaviour that otherwise may lead to a speeding-related collision.

Furthermore, the log-rank tests that examine the survival probability equity for two different groups were significant as the *Pr*<0.05 for all the test groups. This means that the survival probabilities for each of the two tested groups were not equal, both proving and matching the previous results of the *KM* graphs.

5.1.4 Collector Sites Only

This section examines the effectiveness of the deployed *MPE* variables to determine whether the MPE is more efficient for collector roadways. Forty-eight collector sites included 134 mainly *PDO* collisions in 2019. Moreover, there were two major and fourteen minor collisions. The deployed *MPE* hours between two crashes varied between 0 and 280 hours; the number of deployed *MPE* visits ranged from 0 to 86. In addition, the ratio between the number of deployed *MPE* hours to visits reached four hours/visit. Tables 13 and 14 a-d and Figures 11 a-c and 12 below show the outputs after executing the methodology.

5.1.4.1 Cox-proportional hazard model

Table 13: *The Hazard Ratio Estimates for the MPE Variables (Collectors Only).*

	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Hours	0.9851	0.00310	0.000	0.9790	0.99119

a) The total number of deployed MPE hours.

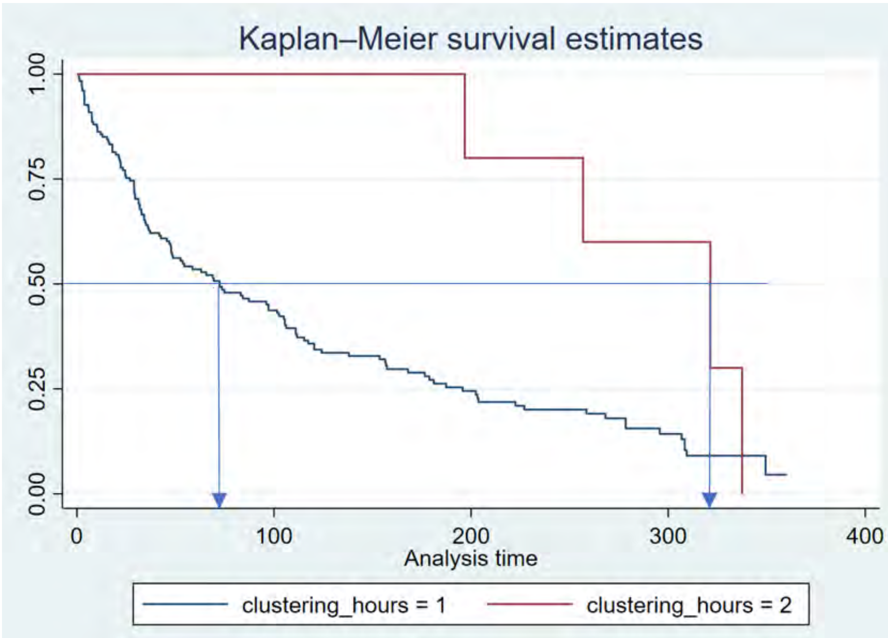
	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Visits	0.9586	0.00809	0.000	0.9429	0.9746

b) The total number of deployed MPE visits.

	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Hours per Visits	0.65281	0.04603	0.000	0.5685	0.7495

c) The ratio between the number of *MPE* visits and hours (*HpV*).

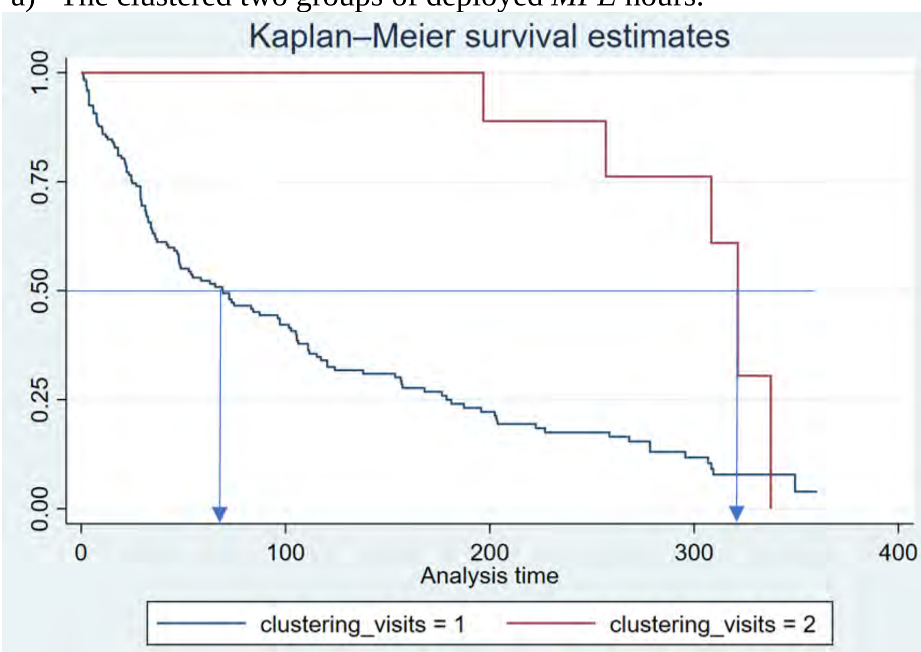
5.1.4.2 Kaplan-Meier Graphs



1 = Below-average

2 = Above-average

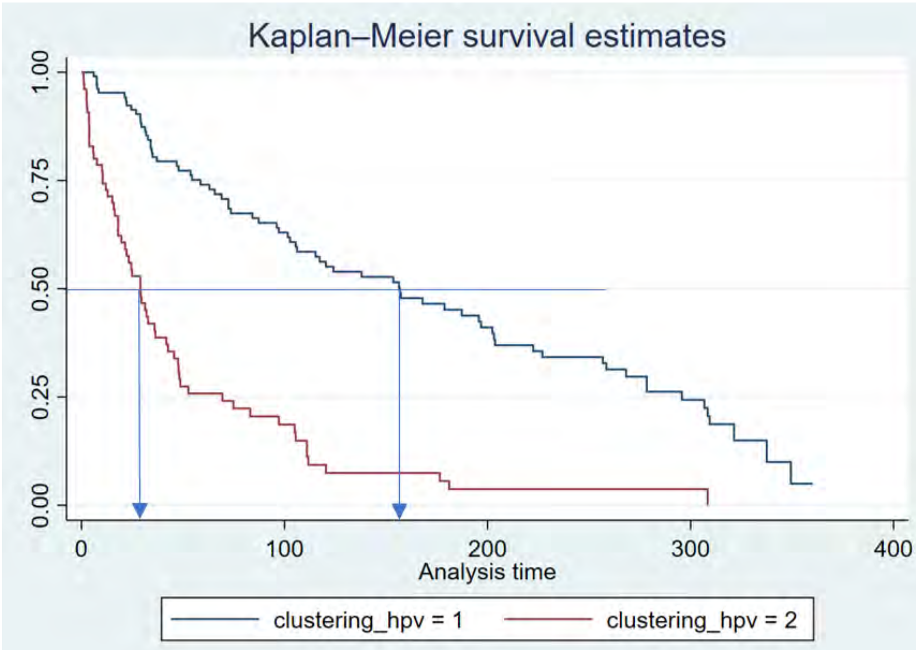
a) The clustered two groups of deployed *MPE* hours.



1 = Below-average

2 = Above-average

b) The clustered two groups of deployed MPE visits.

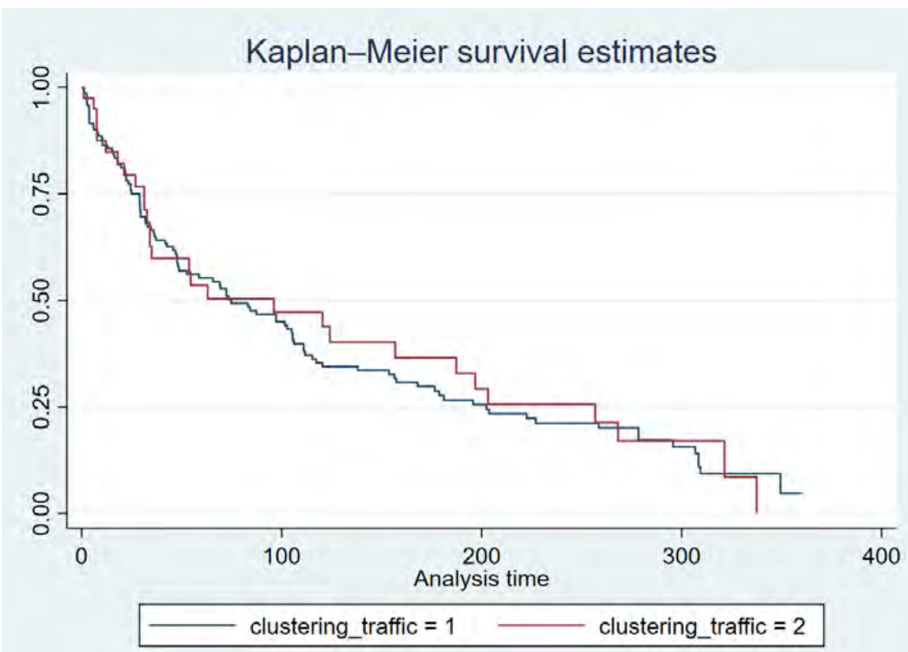


1 = Above-average

2 = Below-average

c) The clustered two groups of deployed *MPE* visits.

Figure 11: *The KM Survival Estimates for the MPE Groups (Collectors Only).*



1 = Below-average

2 = Above-average

Figure 12: *The KM Survival Estimates for the Clustered Traffic Groups (Collectors Only)*

5.1.4.3 Log-rank Equality Test

Table 14: *The Log-Rank Test for Different Groups (Collectors Only).*

Hours	Observed Events	Expected Events
1	130	124.06
2	4	9.94
Total	134	134.00
		Chi2(1) = 4.08 Pr>chi2 = 0.0433

a) Log-rank test for *MPE* hours groups.

Visits	Observed Events	Expected Events
1	129	116.85
2	5	17.15
Total	134	134.00
		Chi2(1) = 10.68 Pr>chi2 = 0.0011

b) Log-rank test for *MPE* visits groups.

HpV	Observed Events	Expected Events
1	71	105.70
2	63	28.30
Total	134	134.00
		Chi2(1) = 58.78 Pr>chi2 = 0.0000

c) Log-rank test for *MPE* hours/visit groups.

Traffic Count	Observed Events	Expected Events
1	105	103.29
2	29	30.71
Total	134	134.00
		Chi2(1) = 0.13 Pr>chi2 = 0.7233

d) Log-rank test for traffic volume groups.

The collectors' site data analysis showed that the *MPE* variables' impacts on the duration between collisions are more beneficial than other categories (i.e., all sites, arterials and collectors, and arterials only groups). For instance, as shown in Figure 11a, the deployed *MPE* hours/visit variable had *HR* = 0.65. This means the consequence of accounting for the *MPE* hours/visit is a reduction in the risk of a collision occurrence by 35%. Also, the deployed *MPE* visits trigger a reduction of 5% in collisions.

The *KM* graphs results provided the same summary as previously explained for different categories. For example, in Figure 11c, the median survival probability (at 0.5 on Y-axis) for above-average *MPE* hours/visit locations is 265 days, compared to 34 days for the sites in the below-average group. Moreover, it is noticeable that the *KM* graph for the traffic volume clusters intersects at many points and yields almost the same survival probability over the year. This indicates that the traffic volume does not impact the survival probability for the collectors' sites. In other words, the traffic volume is not a factor that affects the survival probability of collector roadways.

The log-rank tests were not significant for all groups (i.e., *Pr* value> 0.05 in one case). For instance, in Tables 14 a-c, the *Pr* < 0.05 for the *MPE* hours, visits, and hours/visit groups indicate that these groups have different survival probabilities. As expected, the *Pr* value is greater than 0.05 for the traffic volume groups, which yields the same results as the *KM* graphs. This indicates that the null hypothesis was not rejected, and both traffic volume clusters have similar survival probability.

5.1.5 High Traffic Volume Locations Only

After grouping all sites based on roadway type to study the impact of *MPE* variables on different roadway categories, we classified the sites based on the traffic volume. From this, we compared the effectiveness of deployed *MPE* variables on two different traffic volume categories (i.e., above-average traffic volume sites and below-average traffic volume sites). Ten high-traffic volume sites experienced 133 collisions in 2019. These sites were all arterial roadways. The figures below explain the results:

5.1.5.1 Cox proportional hazard model

Table 15: *The Hazard Ratio Estimates for the MPE Variables (High Traffic Volume Sites Only).*

	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Hours	0.97506	0.00286	0.000	0.96945	0.9806

(a) The total number of deployed *MPE* hours.

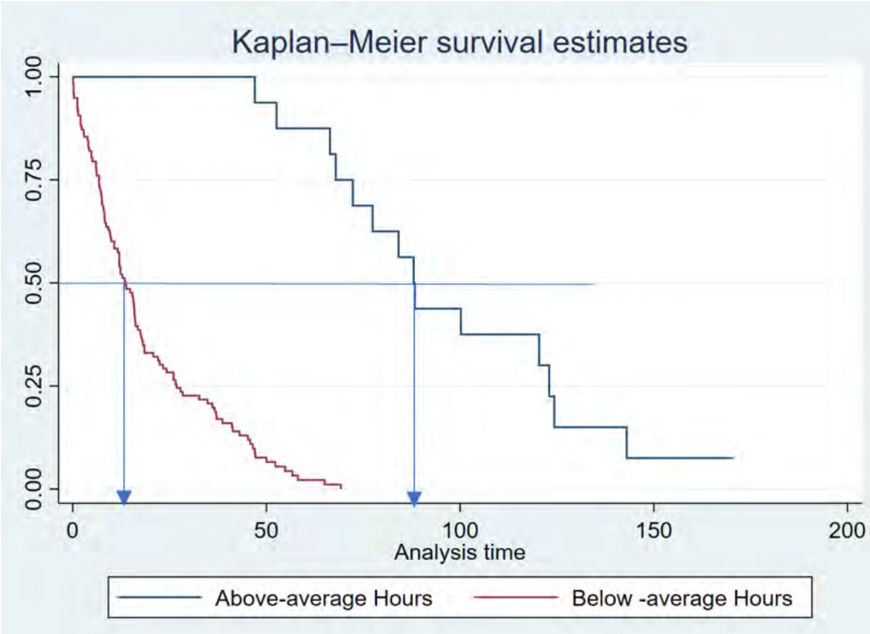
	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Visits	0.9168	0.00908	0.000	0.8991	0.9347

(b) The total number of deployed *MPE* visits.

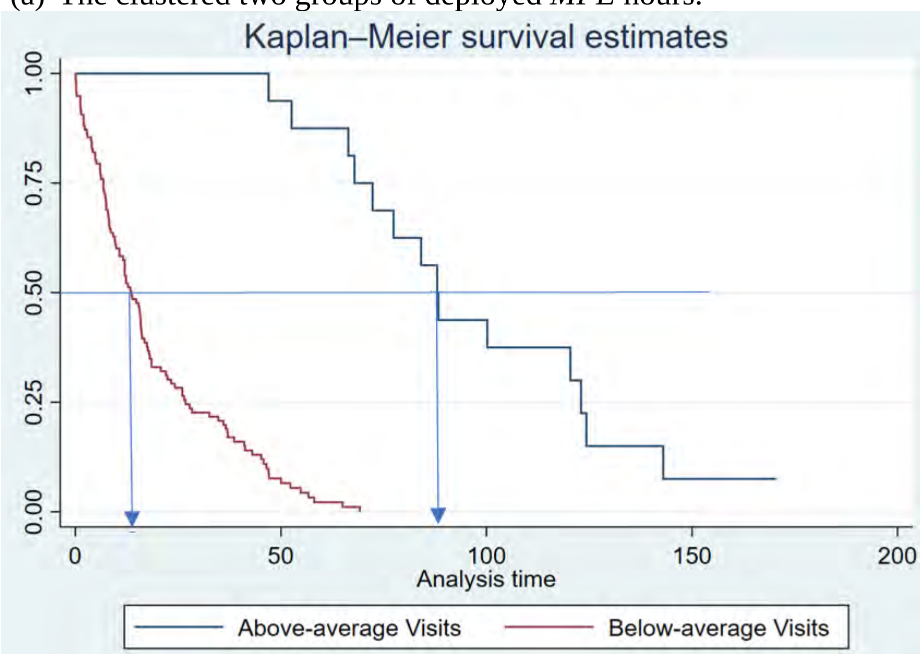
	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Hours per Visits	0.4884	0.0545	0.000	0.3924	0.6078

(c) The ratio between the number of visits and hours (*HpV*).

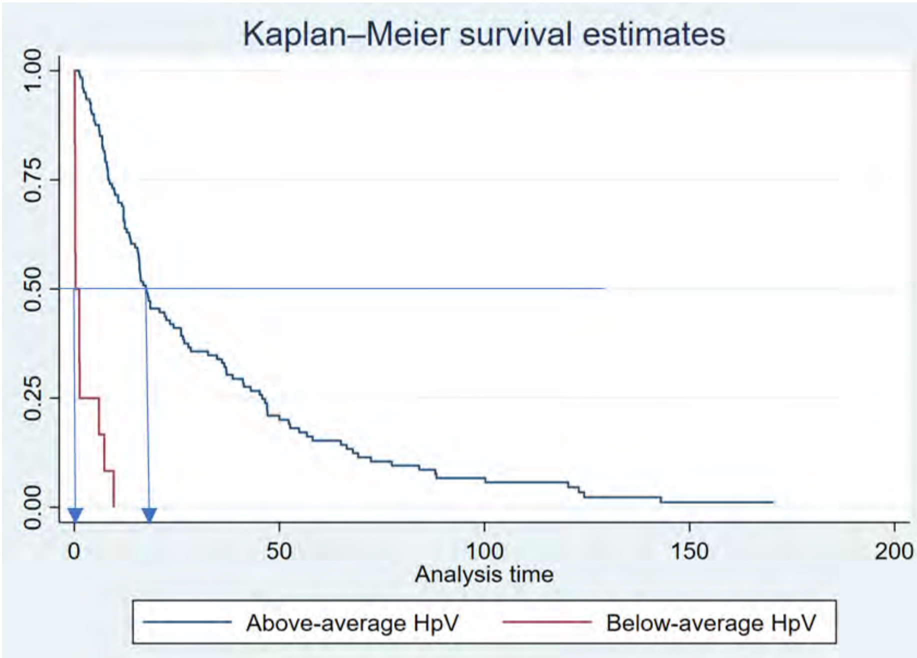
5.1.5.2 Kaplan-Meier Graphs



(a) The clustered two groups of deployed *MPE* hours.



(b) The clustered two groups of deployed *MPE* visits.



(c) The clustered two groups of deployed *MPE* visits.

Figure 13: *The KM Survival Estimates for the MPE Groups (High Traffic Volume Sites Only).*

5.1.5.3 Log-rank Equality Test

***Table 16**: The Log Rank Test for Different Groups (High Traffic Volume Sites Only).*

Hours	Observed Events	Expected Events
Above average Below-average	14 109	44.25 78.75
Total	123	123.00
		Chi2(1) = 48.42 Pr>chi2 = 0.0000

a) Log-rank test for *MPE* hours groups.

Visits	Observed Events	Expected Events
Above average Below-average	14 109	44.25 78.75
Total	123	123.00
		Chi2(1) = 48.42 Pr>chi2 = 0.0000

b) Log-rank test for *MPE* visits groups.

HpV	Observed Events	Expected Events
Above average Below-average	111 12	121.69 1.31
Total	123	123.00
		Chi2(1) = 92.06 Pr>chi2 = 0.0000

c) Log-rank test for *MPE* hours/visit groups.

The procedure's outputs for the high traffic volume sites indicated that the different MPE variables impacted the duration between two consequent collisions. As shown in Table 15, all the *MPE* variables significantly correlate to the time between collisions (95% confidence interval). Table 15c illustrates the deployed MPE hours to visits (HpV) ratio, showing a remarkable influence on the collision occurrence. The hazard ratio for this variable is 0.48, indicating a 52% reduction in the risk of a collision. Moreover, the deployed *MPE* hours and visits had an *HR* of 0.97 and 0.91, respectively. This reflects a decrease in collision occurrence by 3% and 9% by implementing *MPE* hours and visits, respectively. These results comply with the previous conclusion that deploying MPE visits is more effective than deploying *MPE* hours.

The *KM* charts' results showed that the above-average *MPE* sites had a higher survival probability than below-average *MPE* sites. Figure 12a shows that when the survival probability is at 0.5 on Y-axis, the above-average *MPE* hours group is 91 days, while it is 15 days for the below-average group. Similarly, in Figure 12c, the survivability for the above-average *MPE* hours/visit cluster is 21 days on average; the below-average survivability is one day on average. These statistics demonstrate that higher *MPE* hours, visits, or hours/visits are linked to higher survivability and lower risk of collision occurrence, along with the longer durations between two consequent collisions.

The equity tests were significant for all considered groups concerning the log-rank test results. As shown in Table 16, the *Pr* value was less than 0.05, which means the null hypothesis is rejected, and the clusters of each group have different survival probabilities. These results support the *KM* graphs by providing the same conclusion.

5.1.6 Low Traffic Volume Locations

There are 104 low-traffic volume sites that had a total of 450 collisions in 2019. The categories of these sites varied between arterials, collectors, and locals, with a balanced presence of both arterial and collectors' categories. The leading cause of these collisions was following too closely, followed by left turn cross path and stop sign violation. Tables 17 a-c and 18 a-c and Figure 14 a-c below illustrate the results:

5.1.6.1 Cox proportional hazard model

***Table 17**: The Hazard Ratio Estimates for The MPE Variables (Low Traffic Volume Sites Only).*

	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Hours	0.9918	0.00124	0.000	0.9894	0.9943

(a) The total number of deployed *MPE* hours.

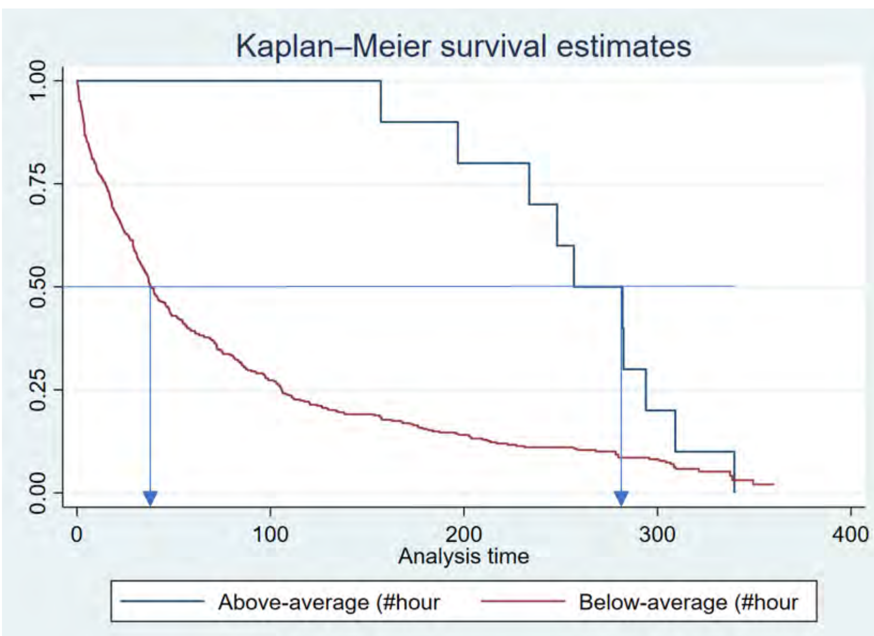
	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Visits	0.9726	0.0037	0.000	0.9652	0.9800

(b) The total number of deployed *MPE* visits.

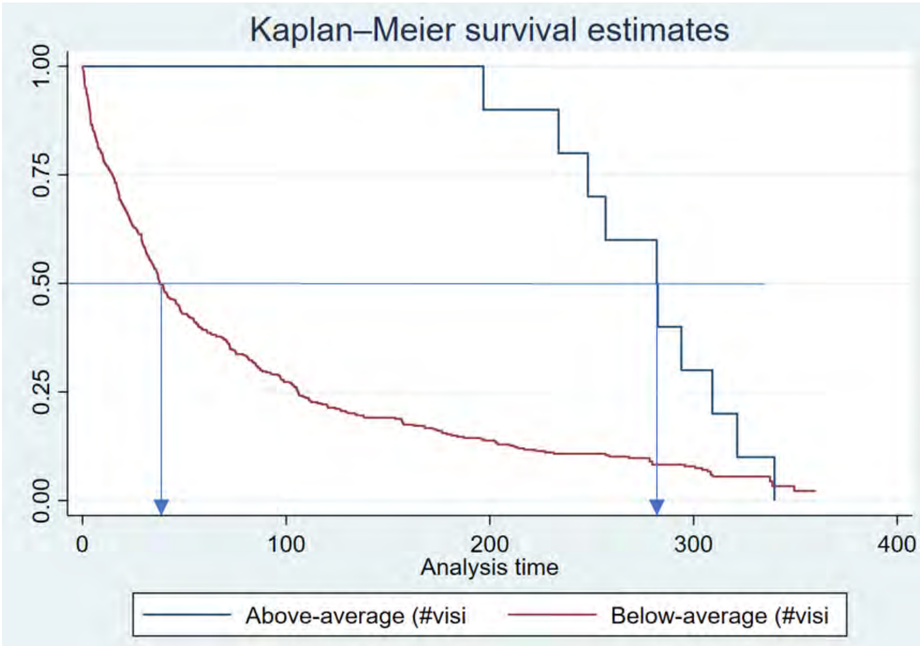
	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Hours per Visits	0.7072	0.0244	0.000	0.6608	0.7569

(c) The ratio between the number of *MPE* visits and hours (*HpV*).

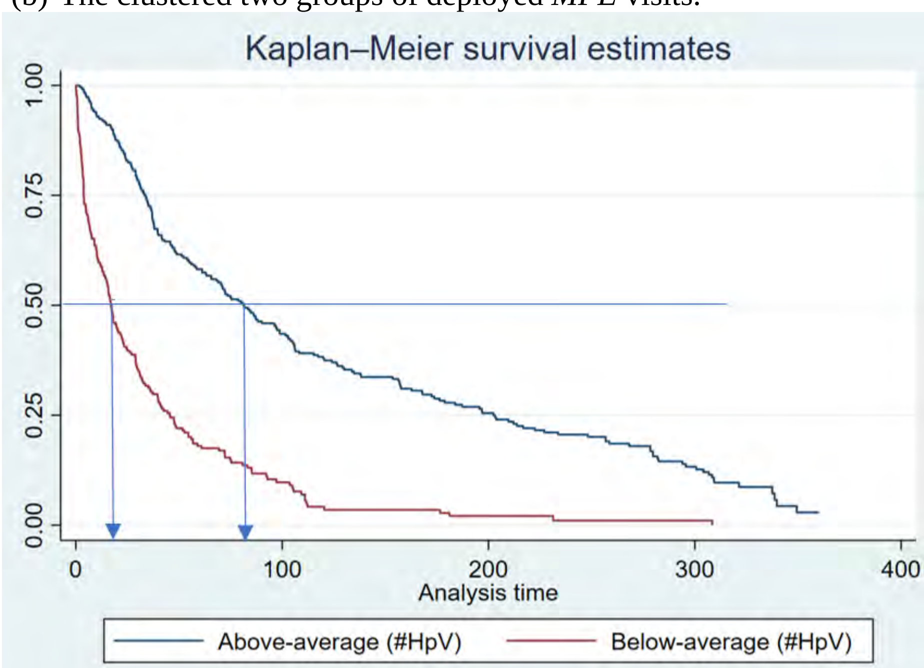
5.1.6.2 Kaplan-Meier Graphs



(a) The clustered two groups of deployed *MPE* hours.



(b) The clustered two groups of deployed *MPE* visits.



(c) The clustered two groups of deployed *MPE* visits.

Figure 14: *The KM Survival Estimates for The MPE Groups (Low Traffic Volume Sites Only)*

5.1.6.3 Log-rank Equality Test

Table 18: *The Log Rank Test for Different Groups (Low Traffic Volume Sites Only).*

Hours	Observed Events	Expected Events
Above average Below-average	10 440	24.78 425.22
Total	450	450.00
		Chi2(1) = 9.67 Pr>chi2 = 0.0019

a) Log-rank test for *MPE* hours groups.

Visits	Observed Events	Expected Events
Above average Below-average	10 440	26.11 423.89
Total	450	450.00
		Chi2(1) = 11.03 Pr>chi2 = 0.0009

b) Log-rank test for *MPE* visits groups.

HpV	Observed Events	Expected Events
Above average Below-average	234 216	340.67 109.33
Total	450	450.00
		Chi2(1) = 149.95 Pr>chi2 = 0.0000

c) Log-rank test for *MPE* hours/visit groups.

The Cox-proportional hazard models for the below-average traffic volume sites showed that the deployed *MPE* hours, visits, and *HpV* significantly reduce the risk of collision occurrence. For instance, as shown in Table 17, the ratio of deployed *MPE* hours to visits has the maximum impact on collisions with an *HR* = 0.7, which means a 30% reduction in the risk of collision occurrence by considering the *MPE HpV*. Moreover, the deployed *MPE* visits have a higher positive impact than MPE hours since it reduces the risk of collision occurrence by 3%.

The clustered groups of MPE hours, visits, and HpV have different survivability for the KM graphs. As presented in Figure 14, the survival probability for the above-average MPE hours is 257 days, compared to forty-five days for the below-average group, and similar survivability for the deployed *MPE* visits groups. Furthermore, for the *HpV MPE* clusters, the above-average sites survived for 138 days on average; on the other hand, the below-average sites survived for twenty-four days.

The log-rank tests were significant for all variable groups. The *Pr* values for *MPE* hours, visits, and *HpV* clusters are zero, emphasizing these groups' different survivability. These results perfectly match the *KM* graphs' outcome.

5.1.7 Speeding Related Collisions Only

The speeding-related collisions happen due to following too closely, running off the road, or striking a parked vehicle. There were 217 speeding-related collisions that took place in 2019. We considered these collisions to study the impact of deployed *MPE* variables on them. Tables 19 a-c and 20 a-d and Figures 15 a-c and 16 illustrate the results:

5.1.7.1 Cox proportional hazard model

Table 19: *The Hazard Ratio Estimates for The MPE Variables (Speeding Related Collisions Only).*

	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Hours	0.9929	0.0015	0.000	0.9900	0.9959

(a) The total number of deployed *MPE* hours.

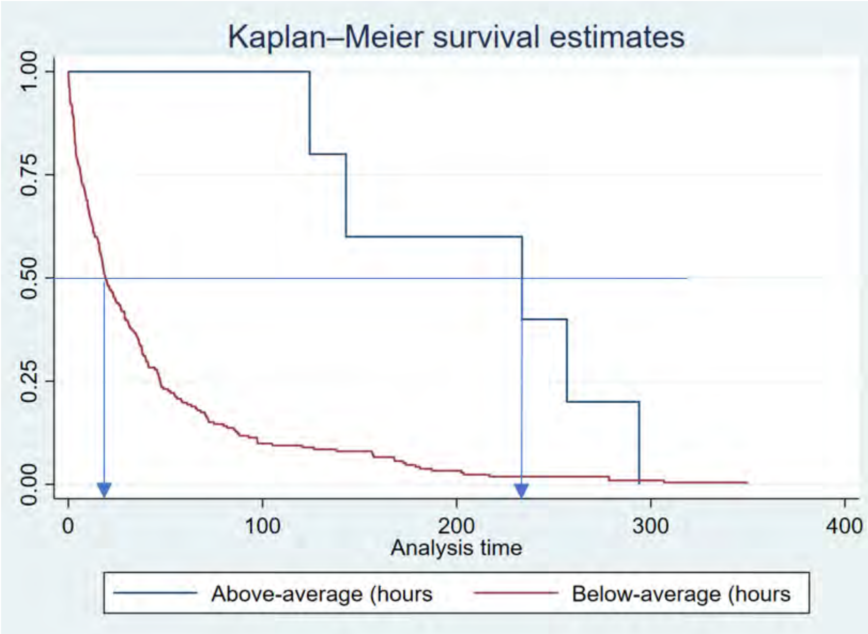
	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Deployment Visits	0.9742	0.0050	0.000	0.9644	0.9841

(b) The total number of deployed *MPE* visits.

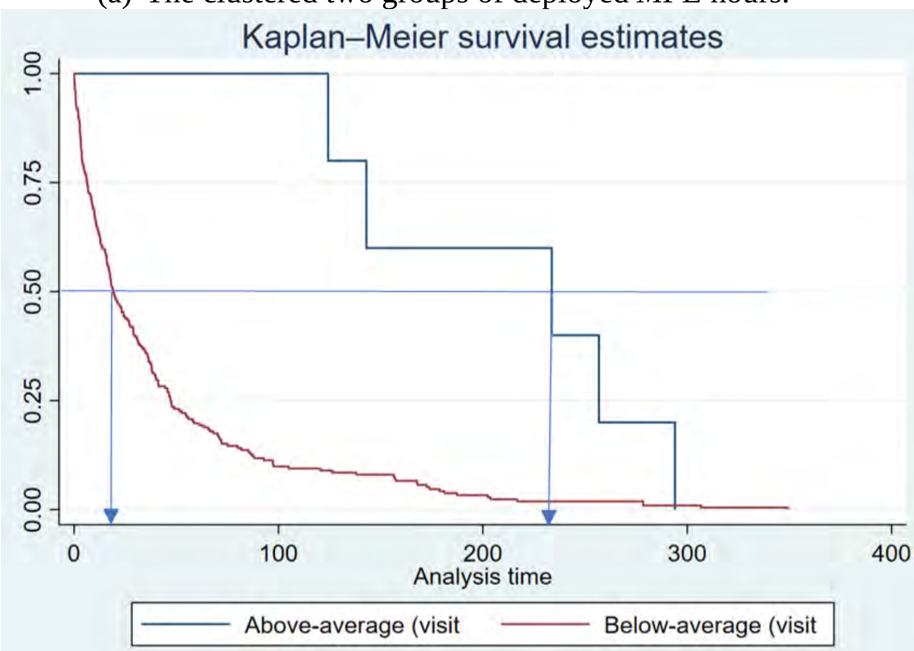
	Haz. Ratio	Std. Err.	P>|z|	[95% Conf. Interval]	
Hours per Visits	0.6789	0.0352	0.000	0.6131	0.7517

(c) The ratio between the number of *MPE* visits and hours (*HpV*).

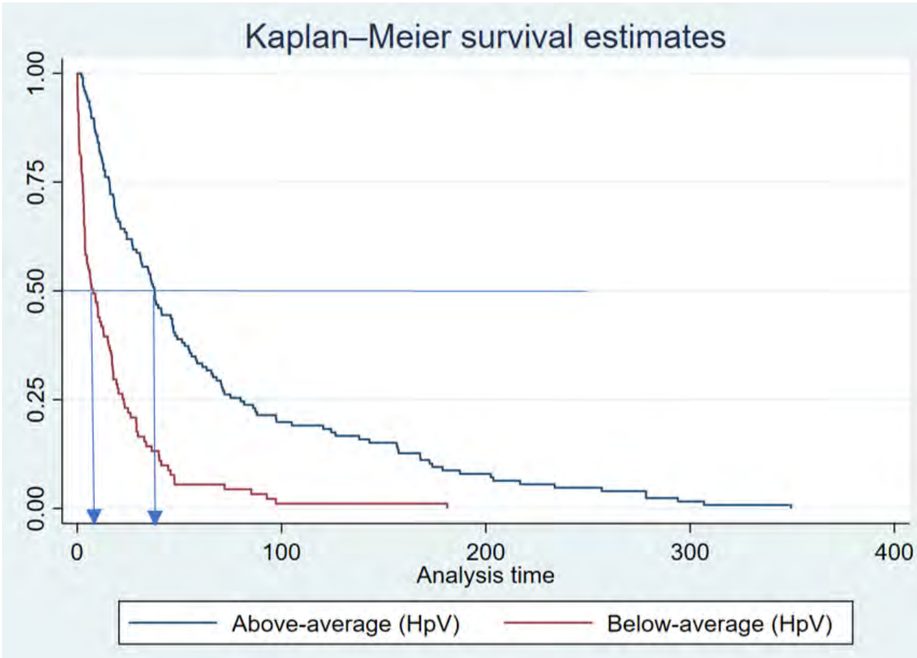
5.1.7.2 Kaplan-Meier Graphs



(a) The clustered two groups of deployed *MPE* hours.



(b) The clustered two groups of deployed *MPE* visits.



(c) The clustered two groups of deployed *MPE* hours/visits.

Figure 15: *The KM Survival Estimates for The MPE Groups (Speeding Related Collisions Only).*

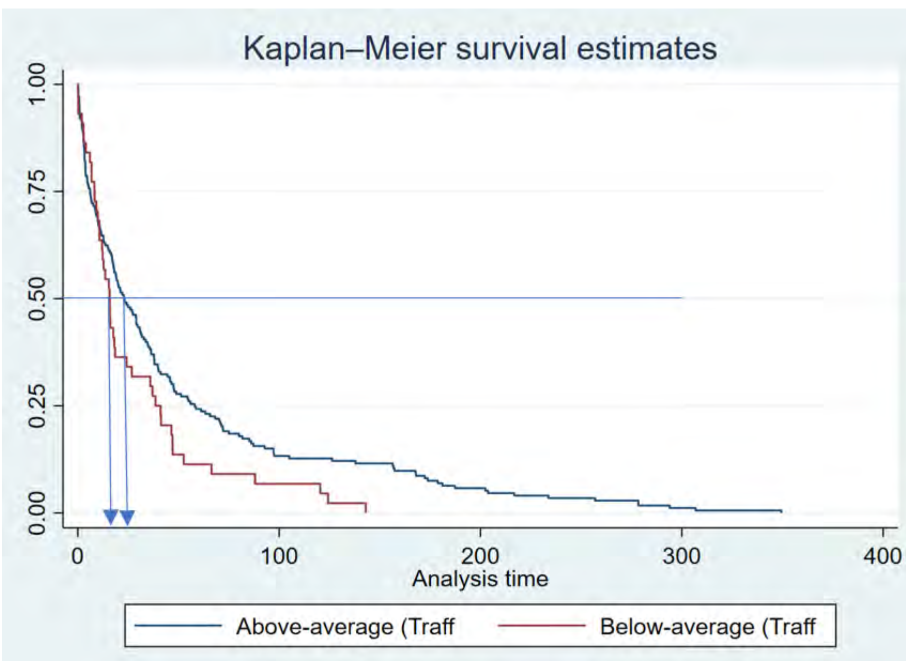


Figure 16: *The KM Survival Estimates for The Clustered Traffic Groups (Speeding Related Collisions Only).*

5.1.7.3 Log-rank Equality Test

Table 20: *The Log Rank Test for Different Groups (Speeding Related Collisions Only).*

Hours	Observed Events	Expected Events
Above average	5	16.16
Below-average	212	200.84
Total	217	217.00
		Chi2(1) = 9.07 Pr>chi2 = 0.0026

a) Log-rank test for *MPE* hours groups.

Visits	Observed Events	Expected Events
Above average	5	16.16
Below-average	12	200.84
Total	217	217.00
		Chi2(1) = 9.07 Pr>chi2 = 0.0026

b) Log-rank test for *MPE* visits groups.

HpV	Observed Events	Expected Events
Above average	126	170.44
Below-average	91	46.56
Total	217	217.00
		Chi2(1) = 58.60 Pr>chi2 = 0.0000

c) Log-rank test for *MPE* hours/visit groups.

Traffic Count	Observed Events	Expected Events
Above average	173	184.06
Below-average	44	32.94
Total	217	217.00
		Chi2(1) = 4.53 Pr>chi2 = 0.0333

d) Log-rank test for traffic volume groups.

These results proved that deploying a higher rate of hours to visits increases the duration between two consequent collisions and hence decreases the probability of collision occurrence. Comparing the results of all site analyses and the speeding-related collisions only shows that the deployed *MPE* is more effective on these specific collision causes. For instance, the *MPE HpV*'s *HR* for the speeding-related collisions is 0.67, which means a 33% reduction in the risk of collision occurrence by accounting for the *MPE HpV* ratio. On the other hand, the *HR* for all types of collisions generally is 22%.

The *KM* graphs in Figures 15a and 15b showed that the average survival probability for above-average *MPE* hours or visits sites is 220 days, while it is seventeen days only for the below-average *MPE* ones. This means the survivability for the sites with above-average *MPE* hours or visits is more than ten times the below-average *MPE* sites. For the deployed *MPE* hours/visit, the survival probability for the below-average site group is ten days, compared to forty-five days for the above-average *MPE* hours/visit sites.

Moreover, the log-rank tests are proof of the *KM* graphs results. As presented in Table 20, the *Pr* values for all clustered groups are less than 0.05, reflecting that these different clusters do not have the same survivability. For instance, the *Pr* equals 0.0026 for the site clusters of *MPE* hours and visits; thus, these results comply with the *KM* graphs.

5.2 Survival Analysis Results (2018, 2017, 2016 and All years)

This methodology was executed separately for each year (2018, 2017, 2016); then, all years were integrated into one study. The results of these years are attached in the appendix section. The results of each year and the combined years yield the same conclusion that the *MPE* variables have a considerable impact on collision occurrence, whereby there is an increase in the duration between two collisions and a corresponding decrease in the risk of collision occurrence. To sum up the results, there is a more significant positive effect from the deployed number of *MPE* visits than the number of deployed *MPE* hours. Moreover, the ratio between *MPE* hours to visits has the most influence on reducing the hazard of collision occurrence. The outcomes of these processes provided similar results to those for the 2019 analysis when applying the analysis procedure to different road categories and traffic volume classifications.

6 CONCLUSION

The main goal of this project has been to develop survival analysis models that investigate the impact of deployed *MPE* hours, visits, and *HpV* on the duration between two consequent collisions in the City of Edmonton. The aim was to mitigate the current issue of speed violators that cause different types of crashes. *MPE* programs are considered an effective solution to restrict irresponsible drivers' behaviours. Therefore, this project explored the survival probability of various locations exposed to different *MPE* hours and visits. Moreover, it detected the potential reduction of the risk of collision occurrence by deploying *MPE* hours and visits.

The proposed methodology consisted of two phases: preparing the provided data and applying the survival analysis. The output of the first step was a Microsoft Excel spreadsheet that contained the necessary survival data. Thus, the outcome table combined the duration between two consequent collisions and the corresponding number of *MPE* hours and visits for each duration. Moreover, the outcome contains collision causes, traffic count, and the first and second collision dates. Next, we categorized and divided the data into groups based on the *MPE* variables and traffic counts. We classified these groups d using *K*-means clusters and employed MATLAB and SAS software to execute this step. The second step was to apply the survival analysis, including Cox proportional hazard models, *KM* survival estimates, and log-rank tests. The Cox proportional hazard models investigated the impact of MPE hours, visits, and *HpV* on the duration between two consequent collisions. In other words, it explored the expected reduction in collision occurrence by deploying *MPE* hours and visits. We plotted the *KM* graphs on different *MPE* clusters to emphasize the influence of the deployed number of MPE hours, visits, and *HpV*. These clusters were classified into two groups (i.e., above average and below average) based on MPE hours, *MPE* visits, *MPE HpV*, and traffic count. Finally, we conducted log-rank tests to explain whether the clusters had the same survival probability, which we expected to comply with the *KM* graphs. All the steps were carried out on different location groups.

The results showed that accounting for the ratio between hours and visits has the maximum impact on increasing the duration between collisions and reducing the risk of collision occurrence. The expected reduction in the collision hazard varied between 52% and 22%, where the maximum reduction could be reached in high traffic volume locations, and the minimum reduction could be expected for all sites. In addition, the deployed *MPE HpV* had a better effect on collector roads compared to arterial roads. We note here that the number of deployed *MPE* visits had a higher impact on increasing the duration between collisions than the number of deployed *MPE* hours. Transportation planners and authorities may wish to consider this for the future to improve the *MPE* program and provide more timing options for critical deployments.

We plotted the *KM* survival estimates for different groups of *MPE* hours, visits, *HpV,* and traffic volume. The graphs' outcomes concluded that the groups of above-average *MPE* hours, *MPE* visits, and *MPE HpV* have higher survivability than the below-average ones. These results emphasize the importance of *MPE* deployment to reduce the risk of collisions occurring. Moreover, the *KM* survival graphs were carried out for traffic count clusters. The results showed that the above-average traffic volume locations had higher survivability than those with below-average traffic volume, possibly because drivers tend to exceed the speed limit when the traffic flow is light. In addition, we carried out log-rank tests. The outputs comply with the conclusions drawn from *KM* survival estimates. Finally, we applied the same methodology to the data from 2016 to 2019.

We recommend exploring the impact of multi-variables in the Cox-proportional hazard models in further research.

REFERENCES

1. Mohan, D., Khayesi, M., Tiwari, G. and Nafukho, F.M. 2006. *Road traffic injury prevention training manual*. World Health Organization.
2. N. H. T. S. A. (2015). Traffic safety facts 2013 data. Available at: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812169 [Accessed May 2022].
3. Transport Canada. (2015). Canadian Motor Vehicle Traffic Collision Statistics 2013. Available at: Canadian Motor Vehicle Traffic Collision Statistics 2013 (canada.ca)
4. Kim, A.M., Wang, X., El-Basyouny, K. and Fu, Q. 2016. Operating a mobile photo radar enforcement program: A framework for site selection, resource allocation, scheduling, and evaluation. *Case Studies on Transport Policy*. 4(3): 218-229. DOI: https://doi.org/10.1016/j.cstp.2016.05.001.
5. Li, R., El-Basyouny, K. and Kim, A. 2019. A city-wide safety analysis of mobile speed enforcement [Presentation]. Available at: https://era.library.ualberta.ca/items/fd7d242b-17d1-456d-840e-a3ec910d2f7c. [Accessed May 2022]. DOI https://doi.org/10.7939/r3-whyz-6832.
6. Alberta Justice and Solicitor General. (2014). Automated Traffic Enforcement Technology Guidelines. Edmonton: Province of Alberta Government. Available at: https://open.alberta.ca/publications/automated-traffic-enforcement-technology-guidelines-2014. [Accessed May 2022].
7. Li, Y. 2019. Planning a Mobile Photo Enforcement Program by Mapping Program Goals to Deployment Decisions. Available at: https://era.library.ualberta.ca/items/4d4cb2a3-55e8-456c-8713-f60949ee584f. [Accessed May 2022]. DOI https://doi.org/10.7939/r3-v4fe-k689.
8. Li, R., El-Basyouny, K., Kim, A. and Gargoum, S. 2017. Relationship between road safety and mobile photo enforcement performance indicators: A case study of the city of Edmonton. *Journal of Transportation Safety & Security*. 9(2): 195-215. DOI https://doi.org/10.1080/19439962.2016.1173154.
9. Rodier, C.J., Shaheen, S.A. and Cavanagh, E. 2007. Automated speed enforcement in the US: A review of the literature on benefits and barriers to implementation.

In *Transportation Research Board 87th Annual Meeting. CD-ROM. Washington, DC*. July.

10. Evans, L. 2004. *Traffic Safety*. Science Serving Society. Bloomfield Hills, MI.
11. Elvik, R. 2005. Speed and road safety: Synthesis of evidence from evaluation studies. *Transportation Research Record*. 1908(1): 59-69. DOI https://doi.org/10.1177/0361198105190800108.
12. Helai, H., Chor, C.H. and Haque, M.M. (2008). Severity of driver injury and vehicle damage in traffic crashes at intersections: A Bayesian hierarchical analysis. Accident Analysis and Prevention 40(1): 45–54. DOI https://doi.org/10.1016/j.aap.2007.04.002.
13. Vanlaar, W., Robertson, R. and Marcoux, K., 2011. Evaluation of the Photo Enforcement Safety Program of the City of Winnipeg Final Report. Traffic Injury Research Foundation.
14. Brault, M., Auger, A. and Montégiani, M. 2007. La sécurité aux intersections: Analyse des comportements des conducteurs au feu rouge. In *Présentation à la XVIIe Conférence Canadienne Multidisciplinaire sur la Sécurité Routière à Montréal*. June.
15. Carnis, L. and Blais, E., 2013. An assessment of the safety effects of the French speed camera program. *Accident Analysis & Prevention*. 51: 301-309. DOI https://doi.org/10.1016/j.aap.2012.11.022.
16. Tay, R. & Barros, A. D., 2011. Should traffic enforcement be unpredictable? The case of red light cameras in Edmonton. *Accident Analysis and Prevention*. 43(3): 955–961. DOI https://doi.org/10.1016/j.aap.2010.11.022.
17. Li, R. and Guo, M., 2015. Competing risks analysis on traffic accident duration time. *Journal of Advanced Transportation*. 49(3): 402-415. DOI https://doi.org/10.1002/atr.1279.
18. Alberta Justice and Solicitor General. (2014). Automated Traffic Enforcement Technology Guidelines. Edmonton: Province of Alberta Government. Available at: https://open.alberta.ca/publications/automated-traffic-enforcement-technology-guidelines-2014. [Accessed May 2022].
19. Carnis L. 2011. Automated speed enforcement: What the French experience can teach us. *Journal of Transportation Safety & Security*. 3(1): 15-26. DOI 10.1080/19439962.2010.551450.

20. Bonneson JA and Zimmerman KH. 2004. Effect of yellow-interval timing on the frequency of red-light violations at urban intersections. *Transportation Research Record*. 1865(1): 20-7. DOI https://doi.org/10.3141/1865-04.
21. Li R, El-Basyouny K and Kim A. 2015. Before-and-after empirical Bayes evaluation of automated mobile speed enforcement on urban arterial roads. Transportation Research Record. 2516(1): 44-52. DOI https://doi.org/10.3141/2516-07.
22. Hojati, A.T., Ferreira, L., Washington, S. and Charles, P. 2013. Hazard-based models for freeway traffic incident duration. *Accident Analysis & Prevention*. 52: 171-181. DOI https://doi.org/10.1016/j.aap.2012.12.037.
23. Shi, Y., Zhang, L. and Liu, P. 2015. Survival Analysis of Urban Traffic Incident Duration: A Case Study at Shanghai Expressways. *Journal of Computers*. 26(1): 29-39.
24. Nam, D. and Mannering, F. 1999. An exploratory hazard-based analysis of highway incident duration. *Transportation Research Part A: Policy and Practice*. 34(2): 85-102. DOI: 10.1016/S0965-8564(98)00065-2.
25. Li, R. and Guo, M., 2015. Competing risks analysis on traffic accident duration time. *Journal of Advanced Transportation*. 49(3): 402-415. DOI https://doi.org/10.1002/atr.1279.
26. Yang, X., Gao, Z., Guo, H. and Huan, M. 2012. Survival analysis of car travel time near a bus stop in developing countries. *Science China Technological Sciences*, 55(8): 2355-2361. DOI: https://doi.org/10.1007/s11431-012-4900-z.
27. Lawpoolsri, S., Li, J., and Braver, E.R. 2007. Do Speeding Tickets Reduce the Likelihood of Receiving Subsequent Speeding Tickets? A Longitudinal Study of Speeding Violators in Maryland. *Traffic Injury Prevention*. 8(1): 26-34.
28. Yang, X., Gao, Z., Guo, H., and Huan, M. 2012. Survival analysis of car travel time near a bus stop in developing countries. *Science China Technological Sciences*, 55(8): 2355-2361. DOI: https://doi.org/10.1007/s11431-012-4900-z.
29. Wu J., Zhang S., Singh A.K. and Qin S. 2018. Hazard-Based Model of Mandatory Lane Change Duration. In *CICTP 2017: Transportation Reform and Change—Equity, Inclusiveness, Sharing, and Innovation*. Reston, VA: American Society of Civil Engineers. 805-811.

30. Kang, G. and Fang, S.E. 2011. Applying survival analysis approach to traffic incident duration prediction. In *ICTIS 2011: Multimodal Approach to Sustained Transportation System Development: Information, Technology, Implementation*. 1523-1531.
31. Hensher, D.A. and Mannering, F.L. 1994. Hazard-based duration models and their application to transport analysis. *Transport Reviews*. 14(1): 63-82. DOI 10.1080/01441649408716866.
32. MacQueen, James. 1967. Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. 1(14).
33. Carroll R. and Lin X. 2008. Nonparametric and semi-parametric regression methods for longitudinal data analysis. *Longitudinal Data Analysis*. London, UK: Chapman & Hall/CRC Handbooks of Modern Statistical Methods. 199-219.
34. Kaplan E.L. and Meier P. 1958. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*. 53(282): 457–481.
35. Rich J.T., Neely J.G., Paniello R.C., Voelker C.C., Nussenbaum B. and Wang E.W. 2010. A practical guide to understanding Kaplan-Meier curves. *Otolaryngology-Head and Neck Surgery*. 143(3): 331-336.
36. Schober, P. and Vetter, T.R. 2018. Survival analysis and interpretation of time-to-event data: The tortoise and the hare. *Anesthesia and Analgesia*. 127(3): 792-798. DOI 10.1213/ANE.0000000000003653
37. Pocock S.J. 1983. *Clinical Trials: A Practical Approach*. New York: Wiley.